BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: Mukerji et al. Examiner: Ramirez, D.M.

Serial No.: 09/624,670 Group Art Unit: 1652

Filed: July 24, 2000

Title: ELONGASE GENES AND USES THEREOF

Case No.: 6407.US.P2

DECLARATION

Assistant Commissioner for
Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

I, Amanda E. Leonard, a United States citizen and resident of Gahanna, Ohio, do declare and say that:

I am familiar with the invention described in the above-referenced application;

That I have read and understand the Office Action of June 18, 2003 and Advisory Action of October 9, 2003 issued in connection with the above-referenced case;

That the data and information presented below establish the necessary support, in terms of enablement, for the subject matter of claims 50-53:

In particular, the Examiner contends in the Office Action and also in the Advisory Action that claims 50-53 remain rejected under 35 U.S.C. 112, first paragraph because the specification, while being enabling for

polynucleotides encoding the polypeptides of SEQ ID NO:63 or SEQ ID NO:64, does not reasonably provide enablement for (1) polynucleotides encoding polypeptides having elongase activity wherein said polypeptides are at least 70% sequence similar to the polypeptides of SEQ ID NO:63 or 64, or (2) polynucleotides encoding polypeptides having elongase activity wherein said polypeptides are at least 60% sequence identical to the polypeptides of SEQ ID NO:63 or 64. In particular, the Examiner notes that the specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention commensurate in scope with the claims.

In response, it is submitted that known elongase amino acid sequences have one highly conserved motif, a histidine box, and many invariant residues that are identical (shaded boxes in Figure 4 of reference 1). These conserved residues suggest their critical importance in the catalytic activity and structure of the elongases (1, 2). Alterations of these conserved residues have been shown to cause a reduction in enzyme activity or enzyme inactivity, (3, 4). There are approximately 130 amino acids between the first conserved residue (lysine; K) and the last conserved residue (phenylalanine; F). There is a

relatively high degree of conservation over this limited 'core region' which is required for elongase activity; however, there is very little conservation over the large stretches comprising the rest of the sequences (1,5).

Of this core region, an elongase must have the amino acid sequence of $KXXEXXDTX_{(4)}LX_{(11-13)}HHX_{(22-25)}NX_{(3)}HXXMYXYY$. More specifically, the variant histidine box of *Isochrysis galbana* elongase, IgASE1, was shown to be required for optimum elongase activity (3). The histidine box is contained within approximately 28 amino acids of the core region (H-box region hereafter).

MELO4 and MELO7 amino acid sequences have 58.9% identity overall, 73.8% identity in the core region, and 71.4% identity in the H-box region (see Figure 1, attached). The identities of MELO4 and MELO7 H-box regions to those of active elongases are shown in Table I. The H-box region alignments are shown in Figures 2 (attached) and 3 (attached).

Table I. Identities of active elongase H-box regions.
NS = no significant identity

	MELO4	MELO7
ELO1	53%	42%
ELO2	57%	53%
ELO3	57%	42%
ELOVL1	60%	64%
ELOVL2	92%	67%
ELOVL3	46%	NS

ELOVL4	75%	64%
ELOVL5	71%	100%
LCE	53%	50%
ELG4	75%	64%
Elovl1	60%	64%
Elovl2	MELO4	71%
Elovl3	50%	NS
Elovl4	75%	64%
Lce	53%	50%
rELO1	71%	100%
rELO2	53%	50%
CEELO1	60%	50%
GLELO	60%	60%
MAELO	50%	39%
IgASE1	46%	28%
PSE1	60%	57%
MELO7	71%	MELO7

Therefore, it can be concluded that an elongase, such as those encoded by the polynucleotides of claims 50-53, must have a histidine box and the invariant amino acid sequences within the core region, and it must demonstrate elongase activity by adding two carbon units to the carboxyl end of a fatty acid chain. This elongase must comprise an amino acid sequence having at least 60% identity to the H-box region (28 amino acid sequence of the core region) of the amino acid sequences set forth in SEQ ID NO:63 and SEQ ID NO:64.

In view of the above, there is certainly a correlation between the H-box region of most elongases and the function of the elongase. Thus, there is a correlation between structure and function. In particular, if the H-box region of an elongase possesses at least 60% amino acid sequence

identity to the H-box region of SEQ ID NO:63 and SEQ ID NO:64, the elongase will function properly. Thus, in view of the references presented, that which is currently known in the art and alignment data presented above, claims 50-53 are fully supported.

Furthermore, the undersigned declarant declares that all statements made herein of her own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Amanda Erin-Yeong Leonard
Amanda E. Leonard

1/14/04
date

Summary of Attachments to Declaration of A. Leonard:

1) **Figure 1** shows the alignments of MELO4 and MELO7.

2) **Figure 2** shows the alignments of MELO4 to known, active elongases disclosed in reference no. 1 (Leonard et al., Progress in Lipid Research 2004; 43:36-54) and reference no. 2 (PCT Application - International Publication Number WO 02/44320 A2).

3) **Figure 3** shows the alignments of MELO7 to known, active elongases disclosed in reference no. 1 ((Leonard et al., Progress in Lipid Research 2004; 43:36-54) and reference no. 2 (PCT Application - International Publication Number WO 02/44320 A2).

4) **References:**

1. Leonard et al., Progress in Lipid Research 2004; 43:36-54 (to be published shortly)
2. PCT Application - International Publication No. WO 02/44320
3. Qi et al., FEBS 2003; 547:137-139
4. Zhang et al., Nature Genetics 2001; 27:89-93
5. Qi et al., FEBS 2002; 510:159-165

```
Smith-Waterman score: 1238; 58.9% identity in 285 aa overlap (4-288:1-282)

                10        20        30        40        50        60
MELO4.pep    MEQLKAFDNEVNAFLDNMFGPRDSRVRGWFLLDSYLPTFILTITYLLSIWLGNKYMKNRP
               :: ||  :::::  ::||||:||:|||||||:|:|||: :: ||| :||| ||||||
MELO7.pep      MEHFDASLSTYFKAFLGPRDTRVKGWFLLDNYIPTFVCSVIYLLIVWLGPKYMKNRQ
                       10        20        30        40        50

                70        80        90       100       110       120
MELO4.pep    ALSLRGILTLYNLAITLLSAYMLVELILSSWEGGYNLQCQNLDSAGEGDVRVAKVLWWYY
              :| |||| ||||::|||| ||: ||: : ||| ||: ||:  ||||:|::: :||||||
MELO7.pep    PFSCRGILQLYNLGLTLLSLYMFYELVTGVWEGKYNFFCQGTRSAGESDMKIIRVLWWYY
              60        70        80        90       100       110

               130       140       150       160       170       180
MELO4.pep    FSKLVEFLDTIFFVLRKKTNQITFLHVYHHASMFNIWWCVLNWIPCGQSFFGPTLNSFIH
             ||||:||:||:||:|||:::||| |||||||:|:|||| |:||:|||:|:|| |||||||
MELO7.pep    FSKLIEFMDTFFFILRKNNHQITVLHVYHHATMLNIWWFVMNWVPCGHSYFGATLNSFIH
             120       130       140       150       160       170

               190       200       210       220       230       240
MELO4.pep    ILMYSYYGLSVFPSMHKYLWWKKYLTQAQLVQFVLTITHTLSAVVKPCGFPFGCLIFQSS
             :|||||||||| :|||: |||||||:||:|||||||||| :|  :|  ||:||:| |:|| :
MELO7.pep    VLMYSYYGLSSIPSMRPYLWWKKYITQGQLVQFVLTIIQTTCGVFWPCSFPLGWLFFQIG
             180       190       200       210       220       230

               250       260       270       280       290
MELO4.pep    YMMTLVILFLNFYIQTYRKKPVKKELQEKEVKNGFPKAHLIVANGMTDKKAQ
             ||::|: || ||||||| || ::::   |:  :|  :: : ::|| |::
MELO7.pep    YMISLIALFTNFYIQTYNKKGASRR---KDHLKGHQNGSVAAVNGHTNSFPSLENSVKPR
             240       250       260          270       280       290

MELO7.pep    KQRKDX
                300


Smith-Waterman score: 383;  73.8% identity in 65 aa overlap(123-187:120-184)

                130       140       150       160       170       180
MELO4.pep    KLVEFLDTIFFVLRKKTNQITFLHVYHHASMFNIWWCVLNWIPCGQSFFGPTLNSFIHILMYSYY
             ||:||:||:||:|||:::||| |||||||:|:|||| |:||:|||:|:|| ||||||||:||||||
MELO7.pep    KLIEFMDTFFFILRKNNHQITVLHVYHHATMLNIWWFVMNWVPCGHSYFGATLNSFIHVLMYSYY
           120       130       140       150       160       170       180


Smith-Waterman score: 148; 71.4% identity in 28 aa overlap(123-150:120-147)

                130       140       150
MELO4.pep    KLVEFLDTIFFVLRKKTNQITFLHVYHH
             ||:||:||:||:|||:::||| ||||||
MELO7.pep    KLIEFMDTFFFILRKNNHQITVLHVYHH
           120       130       140
```

Figure 1

```
 Identities = 28/28 (100%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            KLVEFLDTIFFVLRKKTNQITFLHVYHH
Elovl2: 123 KLVEFLDTIFFVLRKKTNQITFLHVYHH 150


 Identities = 26/28 (92%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K VEFLDTIFFVLRKKT+QITFLHVYHH
ELOVL2: 123 KSVEFLDTIFFVLRKKTSQITFLHVYHH 150


 Identities = 21/28 (75%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K VE+LDT+FF+LRKK NQ++FLHVYHH
Elovl4: 135 KGVEYLDTVFFILRKKNNQVSFLHVYHH 162


 Identities = 21/28 (75%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K VE+LDT+FF+LRKK NQ++FLHVYHH
ELOVL4: 135 KGVEYLDTVFFILRKKNNQVSFLHVYHH 162


 Identities = 21/28 (75%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K +E LDTIFFVLRKK +Q+TFLHV+HH
ELG4:   124 KFIELLDTIFFVLRKKNSQVTFLHVFHH 151


Identities = 20/28 (71%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            KL+EF+DT FF+LRK  +QIT LHVYHH
MELO7:  437 KLIEFMDTFFFILRKNNHQITVLHVYHH 464


 Identities = 20/28 (71%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            KL+EF+DT FF+LRK  +QIT LHVYHH
rELO1:  120 KLIEFMDTFFFILRKNNHQITVLHVYHH 147


 Identities = 20/28 (71%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            KL+EF+DT FF+LRK  +QIT LHVYHH
ELOVL5: 120 KLIEFMDTFFFILRKNNHQITVLHVYHH 147


 Identities = 17/28 (60%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K++E +DT+ F+LRKK  Q+TFLHV+HH
Elovl1: 118 KVIELMDTVIFILRKKDGQVTFLHVFHH 145


 Identities = 17/28 (60%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K +E +DT+ F+LRKK  Q+TFLHV+HH
ELOVL1: 118 KFIELMDTVIFILRKKDGQVTFLHVFHH 145
```

Figure 2

```
 Identities = 17/28 (60%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K VEF+DT+  +L++ T QI+FLHVYHH
PSE1:   137 KYVEFMDTVIMILKRSTRQISFLHVYHH 164


 Identities = 17/28 (60%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K++EF+DT+  VL+K   QI+FLHVYHH
GLELO:  161 KIMEFVDTMIMVLKKNNRQISFLHVYHH 188


 Identities = 17/28 (60%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            KL E +DTIF VLRK+   + FLH YHH
CEELO1: 134 KLFELVDTIFLVLRKR--PLMFLHWYHH 159


 Identities = 16/28 (57%)
MELO4:  1   KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K +EF+DT F VL+ K  ++TFLH YHH
ELO2:   157 KFIEFIDTFFLVLKHK--KLTFLHTYHH 182


 Identities = 16/28 (57%)
MELO4: 1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K VE +DT+F VLR+K  ++ FLH YHH
ELO3:  164  KFVELIDTVFLVLRRK--KLLFLHTYHH 189


 Identities = 15/28 (53%)
MELO4: 1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K VEF DT+  VL+ +  ++TFLH YHH
ELO1:  151  KFVEFADTVLMVLKHR--KLTFLHTYHH 176


 Identities = 15/28 (53%)
MELO4: 1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K  E  DTIF +LRK+  ++ FLH YHH
rELO2: 120  KAPELGDTIFIILRKQ--KLIFLHWYHH 145


 Identities = 15/28 (53%)
MELO4: 1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K  E  DTIF +LRK+  ++ FLH YHH
Lce:   120  KAPELGDTIFIILRKQ--KLIFLHWYHH 145


 Identities = 15/28 (53%)
MELO4: 1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K  E  DTIF +LRK+  ++ FLH YHH
LCE:   120  KAPELGDTIFIILRKQ--KLIFLHWYHH 145


 Identities = 14/28 (50%)
MELO4: 1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K+VE  DT F +LRK+   + F+H YHH
Elovl3:125  KVVELGDTAFIILRKRP--LIFVHWYHH 150
```

Figure 2 cont'd

```
 Identities = 14/28 (50%)
MELO4: 1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K  E  DT+F VL+KK   + FLH +HH
MAELO: 141  KYWELADTVFLVLKKK--PLEFLHYFHH 166


 Identities = 13/28 (46%)
MELO4: 1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K VE+LDT + VL+ K  +++FL  +HH
IgASE1:351  KYVEYLDTAWLVLKGK--RVSFLQAFHH 416


 Identities = 13/28 (46%)
MELO4: 1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28
            K++E  DT F +LRK+   + F+H YHH
ELOVL3:124  KVIELGDTAFIILRKR--PLIFIHWYHH 149
```

Figure 2 cont'd

```
 Identities = 28/28 (100%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             KLIEFMDTFFFILRKNNHQITVLHVYHH
rELO1:  120  KLIEFMDTFFFILRKNNHQITVLHVYHH 147


 Identities = 28/28 (100%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             KLIEFMDTFFFILRKNNHQITVLHVYHH
ELOVL5: 120  KLIEFMDTFFFILRKNNHQITVLHVYHH 147


 Identities = 20/28 (71%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             KL+EF+DT FF+LRK  +QIT LHVYHH
MELO4:  1    KLVEFLDTIFFVLRKKTNQITFLHVYHH 28


 Identities = 19/28 (67%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K +EF+DT FF+LRK   QIT LHVYHH
ELOVL2: 123  KSVEFLDTIFFVLRKKTSQITFLHVYHH 150


 Identities = 18/28 (64%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K +E++DT FFILRK N+Q++ LHVYHH
Elovl4: 135  KGVEYLDTVFFILRKKNNQVSFLHVYHH 162


 Identities = 18/28 (64%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K +E++DT FFILRK N+Q++ LHVYHH
ELOVL4: 135  KGVEYLDTVFFILRKKNNQVSFLHVYHH 162


 Identities = 18/28 (64%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K+IE MDT  FILRK + Q+T LHV+HH
Elovl1: 118  KVIELMDTVIFILRKKDGQVTFLHVFHH 145


 Identities = 18/28 (64%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K IE MDT  FILRK + Q+T LHV+HH
ELOVL1: 118  KFIELMDTVIFILRKKDGQVTFLHVFHH 145


 Identities = 18/28 (64%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K IE +DT FF+LRK N Q+T LHV+HH
ELG4:   124  KFIELLDTIFFVLRKKNSQVTFLHVFHH 151


 Identities = 17/28 (60%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K++EF+DT   +L+KNN QI+ LHVYHH
GLELO:  161  KIMEFVDTMIMVLKKNNRQISFLHVYHH 188
```

Figure 3

```
 Identities = 16/28 (57%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K +EFMDT   IL+++  QI+ LHVYHH
PSE1:   137  KYVEFMDTVIMILKRSTRQISFLHVYHH 164


 Identities = 15/28 (53%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K IEF+DTFF +L+  + ++T LH YHH
ELO2:   157  KFIEFIDTFFLVLK--HKKLTFLHTYHH 182


 Identities = 14/28 (50%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K  E  DT F ILRK   ++  LH YHH
Lce:    120  KAPELGDTIFIILRK--QKLIFLHWYHH 145


 Identities = 14/28 (50%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K  E  DT F ILRK   ++  LH YHH
LCE:    120  KAPELGDTIFIILRK--QKLIFLHWYHH 145


 Identities = 14/28 (50%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K  E  DT F ILRK   ++  LH YHH
rELO2:  120  KAPELGDTIFIILRK--QKLIFLHWYHH 145


 Identities = 14/28 (50%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             KL E +DT F +LRK    +  LH YHH
CEELO1: 134  KLFELVDTIFLVLRK--RPLMFLHWYHH 159


 Identities = 12/28 (42%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K +E +DT F +LR+   ++  LH YHH
ELO3:   164  KFVELIDTVFLVLRRK--KLLFLHTYHH 189


 Identities = 12/28 (42%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K +EF DT   +L+  + ++T LH YHH
ELO1:   151  KFVEFADTVLMVLK--HRKLTFLHTYHH 176


 Identities = 11/28 (39%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K  E  DT F +L+K    +  LH +HH
MAELO:  141  KYWELADTVFLVLKKK--PLEFLHYFHH 166


 Identities = 8/28 (28%)
MELO7:  1    KLIEFMDTFFFILRKNNHQITVLHVYHH 28
             K +E++DT + +L+    +++ L  +HH
IgASE1: 118  KYVEYLDTAWLVLK--GKRVSFLQAFHH 143
```

Figure 3 cont'd

PERGAMON

Progress in Lipid Research 43 (2004) 36–54

Progress in Lipid Research

www.elsevier.com/locate/plipres

Review

Elongation of long-chain fatty acids

Amanda E. Leonard[a], Suzette L. Pereira[a], Howard Sprecher[b], Yung-Sheng Huang[a,*]

[a] *Strategic Research, Ross Products Division, Abbott Laboratories, 625 Cleveland Avenue, Columbus, OH 43215, USA*
[b] *Department of Molecular and Cellular Biochemistry, The Ohio State University, Columbus, OH, USA*

Accepted 13 May 2003

Contents

1. General introduction 36

2. Fatty acid chain elongation 40
   2.1. Introduction 40
   2.2. Elongation in yeast 40
   2.3. Elongation in mammals 41
   2.4. Elongation in plants 43
   2.5. Elongation in lower eukaryotes 45
      2.5.1. Nematode 46
      2.5.2. Fungus 46
      2.5.3. Marine protist 46
      2.5.4. Microalga 47
      2.5.5. Moss 47
   2.6. Sequence alignments/phylogenetic tree 47

3. Conclusion 48

References 50

..... 51

1. General introduction

Polyunsaturated fatty acids (PUFAs) are fatty acids of 18 carbons or more in length with two or more double bonds. Depending on the position of the first double bond proximate to the methyl end of fatty acids, PUFAs can be classified into two major groups, n-6 and n-3 families. They are obtained either through diet or synthesized from dietary essential fatty acids (EFA), i.e.,

* Corresponding author. Tel.: +1-614-624-3501; fax: +1-614-624-0008.
E-mail addresses: vic.haung@abbott.com (Y.-S. Huang).

0163-7827/03/$ - see front matter
doi:10.1016/S0163-7827(03)00040-7

Nomenclature

AA	arachidonic acid
ACP	acyl carrier protein
ADA	adrenic acid
ALA	α-linolenic acid
DHA	docosahexaenoic acid
DPA	docosapentaenoic acid
EFA	essential fatty acids
ELG	human elongase
ELO	yeast fatty acid elongase
ELOVL	human long-chain fatty acid elongase
Elovl	mouse long-chain fatty acid elongase
EPA	eicosapentaenoic acid
EST	expressed sequence tag
ETA	eicosatetraenoic acid
ETrA	eicosatrienoic acid
FACE	fatty acid chain elongation
FAE	fatty acid elongase
FAS	fatty acid synthase
GLA	γ-linolenic acid
GLELO	*Mortierella alpina* γ-linolenic acid elongase
HELO	human fatty acid elongase
IgASE	*isochrysis galbana* elongase
KAS	β-ketoacyl ACP synthase
KCR	β-ketoacyl CoA reductase
KCS	β-ketoacyl CoA synthase
LA	linoleic acid
LCE	human long-chain fatty acid elongase
Lce	mouse long-chain fatty acid elongase
MAELO	*Mortierella alpina* elongase
ORF	open reading frame
PCR	polymerase chain reaction
PSE	*Physcomitrella patens* elongase
PUFA	polyunsaturated fatty acids
rELO	rat fatty acid elongase
SDA	stearidonic acid
SREBP	sterol regulatory element-binding protein
THA	tetracosahexaenoic acid
TPA	tetracosapentaenoic acid
VLCFA	very long-chain fatty acids

linoleic acid (LA, C18:2n-6) and α-linolenic acid (ALA,C18:3n-3), via a common desaturase/elongase system (Fig. 1) [6]. For example, docosahexaenoic acid (DHA, C22:6n-3), a long-chain n-3 PUFA, is synthesized from ALA by the addition of a double bond by a Δ6-desaturase to form stearidonic acid (SDA, C18:4n-3); the elongation of SDA to form eicosatetraenoic acid (ETA, C20:4n-3); the addition of another double bond by a Δ5-desaturase to form eicosapentaenoic acid (EPA, C20:5n-3); the elongation of EPA to form ω3-docosapentaenoic acid (ω3-DPA, C22:5n-3); and the final addition of a double bond to form DHA [104]. The formation of DHA from EPA occurs via two different mechanisms in eukaryotes. In higher eukaryotes like mammals, EPA is elongated to ω3-DPA which is further elongated to ω3-tetracosapentaenoic acid (TPA, C24:5n-3); a double bond is then added to ω3-TPA by a Δ6-desaturase to form ω3-tetracosahexaenoic acid (THA, C24:6n-3), and the THA is oxidized to DHA in peroxisomes [35,107]. However in lower eukaryotes like the *Thraustochytrid* sp. EPA is elongated to ω3-DPA; a double bond is then added directly to ω3-DPA by Δ4-desaturase to generate DHA [87]. To synthesize long chain n-6 PUFAs, ω6-docosapentaenoic acid (ω6-DPA, C22:5n-6) from LA, the

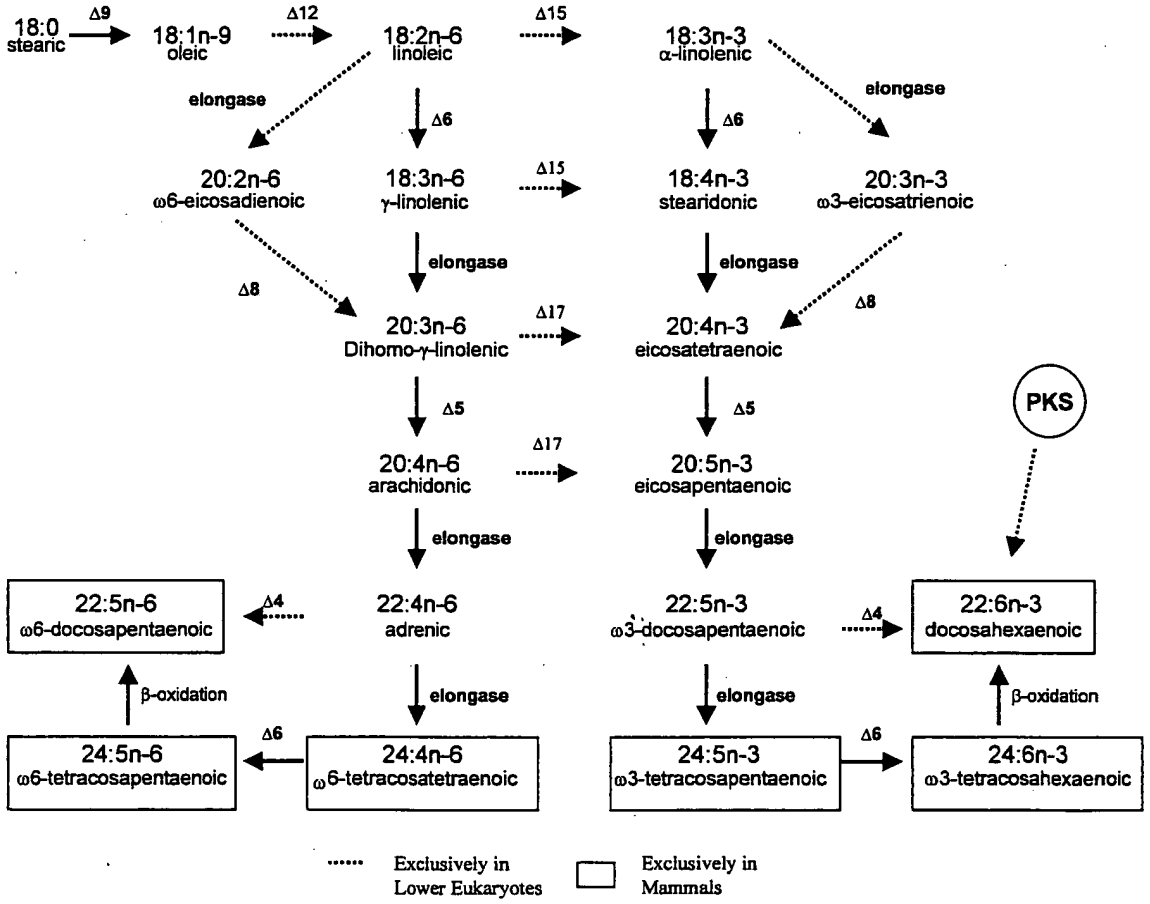


Fig. 1. Biosynthesis of long-chain polyunsaturated fatty acids. Arrows with solid line (→) are found both in mammals are lower eukaryotes, while arrows with dotted line (···►) are exclusively for lower eukaryotes. Fatty acids in □ indicate the pathway is exclusively in mammals.

same alternating desaturation and elongation steps (via either the Δ4-desaturase route, or the β-oxidation route) are utilized.

PUFAs are important structural components that confer membrane fluidity and selective permeability [117]. For example, DHA, a long-chain n-3 PUFA, and arachidonic acid (AA, C20:4n-6), a long-chain n-6 PUFA, are found in high proportions in neuronal tissues such as brain and retina, and testis [12,89]. Deficiencies in both DHA and AA have been associated with disorders of the neuro-visual development and other complications of premature birth [15,26,48,61]. PUFAs also serve as precursors for eicosanoids, growth regulators and hormones, and are constituents of membrane phospholipids involved in signal transduction [53,102], and therefore have profound effects on various physiological processes such as cognitive function, and immunosuppressive and anti-inflammatory actions [100].

Normally, only a very small proportion of dietary linoleate and α-linolenate can be converted to longer PUFAs; most of them are β-oxidized to provide energy [27]. The already low formation of PUFAs can be further depressed by various nutritional and hormonal factors [13]. For example, the levels of PUFAs as compared with healthy subjects, are low in patients suffering from chronic diseases, such as diabetes, and hypercholesterolemia [13]. Riemersma et al. [90] have demonstrated in males that there exists a correlation between the increased mortality from coronary heart disease and low levels of linoleic acid, dihomo-gamma-linolenic acid (C20:3n-6) and arachidonic acid (C20:4n-6) in adipose tissue. Singer et al. [101] have also shown a defective desaturation and elongation in the n-6 and n-3 fatty acid system in patients suffering from hypertension. Imbalance in n-6 and n-3 PUFA metabolism has also been implicated in other chronic diseases such as rheumatoid arthritis, autoimmune disorders, Crohn's disease, and cancer [24,99].

Increasing evidence has shown that dietary supplementation of PUFAs, such as γ-linolenic acid (GLA, C18:3n-6), and EPA/DHA, can exert the anti-inflammatory, anti-thrombotic and anti-arrhythmia activities, and provide beneficial effects on glucose and lipid metabolism [4,17,49,54,58,124,131]. These findings have received much attention from food manufacturers, pharmaceutical companies, as well as the general public. As a result, additions of long-chain PUFAs have been made in new food products such as 'DHA plus' eggs, and DHA-, AA- and/or GLA-fortified infant formulas. The increase in demand has also raised interest in obtaining these PUFAs from other alternate more sustainable sources, particularly by genetic engineering of oil seed crops such as canola and soybean. In recent years, various efforts from many laboratories have been directed towards the identification and cloning of genes encoding the enzymes involved in biosynthesis of PUFAs, for this very use [1].

Since the desaturation steps have long been considered the rate-limiting steps for the biosynthesis of PUFAs, most studies have focused on the isolation and characterization of genes encoding desaturases involved in PUFA production. Genes encoding the following desaturases have been isolated and identified from various sources; a Δ4-desaturase from *Thraustochytrium* sp. [87]; Δ5-desaturases from *Mortierella alpina* [57,69], *Caenorhabditis elgans* [70,121], and human [19,65]; Δ6-desaturases from *M. alpina* [45], cyanobacteria [88,96], borage [88], *C. elegans* [39,73], mammals [2,18,94], and *Physcomitrella patens* [37]; a Δ6/Δ5 fusion desaturase from zebrafish [41]; Δ8-desaturase from *Euglena gracilis* [120]; Δ12 desaturase from cyanobacteria [88], algae [93], *C. elegans* [79], fungi [45,78,92], and plants [14,25,33,80]; and ω-3 fatty acid desaturases from plants [126], cyanobacteria [91], and *C. elegans* [108].

Recently studies in *M. alpina*, however, indicate that the elongation of C18:3n-6 to C20:3n-6 is the rate-limiting step in AA biosynthesis [125]. Thus in this case, increasing the activity of the elongase enzyme via genetic manipulation, rather than manipulating the Δ12-, Δ6-, and Δ5-desaturases for increased enzymatic activity, would be a better approach to increasing the production of AA in this organism. Regardless of which enzyme is responsible for directing the synthesis of PUFAs, it is obvious that elongases are critical for this biosynthetic process. To date, only a few genes involved in the elongation of PUFAs have been characterized. This review covers the recent advances in fatty acid chain elongation in mammals and lower eukaryotes.

2. Fatty acid chain elongation

2.1. Introduction

The terms such as elongase system, elongase, or fatty acid chain elongation system (FACES) all refer to enzymes that are responsible for the addition of two carbon units to the carboxyl end of a fatty acid chain. In both plants and animals, the elongase system, as it is referred to in this review, is composed of four enzymes: a condensing enzyme (β-ketoacyl CoA synthase, KCS), β-ketoacyl CoA reductase (KCR), β-hydroxyacyl CoA dehydrase, and *trans*-2-enoyl CoA reductase (Fig. 2). Fatty acid elongation is initiated by the condensation of malonyl-CoA with a long chain acyl-CoA, yielding a β-ketoacyl-CoA in which the acyl moiety has been elongated by two carbon atoms. The formed β-ketoacyl-CoA is then reduced by β-ketoacyl CoA reductase to β-hydroxyacyl-CoA, which is dehydrated by the dehydrase to an enoyl-CoA, and further reduced by the enoyl reductase to yield the elongated acyl-CoA [21].

In the cell, there exist multiple microsomal elongation systems with different chain length specificity. Early data from several laboratories suggested that there are three separate fatty acid

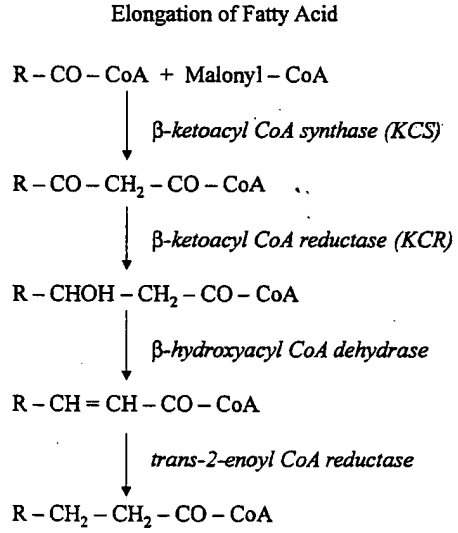


Fig. 2. Fatty acid chain elongation reaction.

elongation systems responsible for the synthesis of saturated fatty acids with different chain lengths: one for the elongation of palmitic acid (C16:0) to stearic acid (C18:0), a second for the elongation of stearic acid to arachidic acid (C20:0), and a third for the elongation of arachidic acid to behenic acid (C22:0) and lignoceric acid (C24:0) [11,38,81]. Suneja et al. [110] have examined the brains of, 'jimpy and quaking' mouse mutants, and found that the chain elongation of arachidoyl CoA (C20:0) and behenoyl CoA (C22:0) were decreased more dramatically than that of palmitoyl CoA (16:0). Additional studies supported the existence of multiple elongation enzymes specific for different chain length saturated fatty acids [111].

There also exist different elongation system responsible for elongation of saturated versus unsaturated fatty acids. Sprecher [103] demonstrated that in animals, fasting followed by re-feeding a normal versus a fat-free diet, differentially modulated the hepatic activities of elongation of palmatic acid versus PUFAs. Further evidence for the existence of different elongation systems for saturated versus unsaturated fatty acids was provided by Bernert and Sprecher [9], Prasad et al. [85], and Chapkin and Coble [16]. More recently, Luthria and Sprecher [66] obtained evidence suggesting the presence of multiple chain-elongating enzymes, specific for fatty acids of different chain length and different degrees of unsaturation.

As mentioned previously, the elongation system requires four enzymatic reactions (Fig. 2). One would question which one of these four enzymes is responsible for the substrate specificity and is rate-limiting. Unfortunately, purification of these membrane-bound elongation enzymes to homogeneity is a daunting task and has thus made the biochemical characterization of these elongation enzymes rather challenging. To date, attempts to purify these enzymes, particularly the condensing enzyme, by many researchers are unsuccessful. Nevertheless, biochemical studies have provided indirect evidence to indicate that the condensing enzyme of the elongation system is rate-limiting [9,21,75]. This enzyme regulates the specificity of the substrate fatty acids in term of chain length and the degree of unsaturation. Its activity is regulated by dietary and hormonal status. Recent advances in genetic manipulation have provided an alternate approach to understand this reaction process. Suneja et al. [110,111] have demonstrated in mouse mutants that a reduction in total elongation activities correlated with the reduction in only the condensation activity, but not the activities of the other three enzymes in the elongation system.

2.2. Elongation in yeast

The elongation products of fatty acid biosynthesis have been much studied, and the organisms that produce high levels of PUFAs had been categorized, but no elongation enzyme has been isolated nor even cloned until recently. Three separate elongase genes, ELO1, ELO2, and ELO3, were identified from the baker's yeast, *Saccharomyces cerevisiae*, and their activities determined by mutation analysis [32,76,97] (Table 1).

The *S. cerevisiae* ELO1 gene product of 310 amino acids was first identified as being responsible for the elongation of C_{14}- fatty acids to C_{16} species, by studying gene defective mutants which are not capable of growing on fatty acids shorter than C16:0 [114]. The ELO2 and ELO3 genes were later identified based on the homology to the ELO1 by searching the *S. cerevisiae* genome database [76]. The yeast ELO2 and ELO3 were initially cloned by complementation of GNS1 mutants and SUR4 mutants, respectively, and categorized as such. Often, a gene was identified, but the activity of the expressed protein was not known at the time. It wasn't until the work of Oh et al.

Table 1
A list of the various members of the ELO family of proteins

Genes	Source	Alternate names	Conversion	Amino acids	Accession number
ELO1	*S. cerevisiae*		C_{14} to C_{16}	310	NP_012339
ELO2	*S. cerevisiae*	GNS1, FEN1	Up to C_{24} SFA/MUFA	347	NP_009963
ELO3	*S. cerevisiae*	SUR4	Broad SFA/MUFA, essential for C_{24} to C_{26}	345	NP_013476
ELOVL1	Human	ELG1	GLA, AA, ALA, EPA	279	XP_002040
ELOVL2	Human	ELG3	C_{20} and C_{22} PUFAs	296	NP_060240
ELOVL3	Human	Cig30, ELG6	GLA, ALA	270	AAG17875
ELOVL4	Human		C_{20} and C_{22} PUFA (?)	314	NP_073563
ELOVL5	Human	HELO1	GLA, AA, STA, EPA, ALA	299	NP_068586
LCE	Human	FACE	C_{12}–C_{16} SFA/MUFA	265	NP_076995
ELG4	Human		GLA, AA, ALA, EPA	281	
Elovl1	Mouse	Ssc1	Broad SFA/MUFA up to C_{24}	279	NP_062295
Elovl2	Mouse	Ssc2	C_{20} and C_{22} PUFAs	292	NP_062296
Elovl3	Mouse	Cig30		271	NP_031729
Elovl4	Mouse			312	AAG47667
Lce	Mouse	FACE	C_{12}–C_{16} SFA/MUFA	267	AY053453
rELO1	Rat		C_{16}–C_{20} MUFA/PUFA	299	NP_599209
rELO2	Rat		C_{16} and C_{18} SFA/MUFA	267	BAB69888
CEELO1	*C. elegans*		C_{16} MUFA/C_{18} PUFA	288	AF244356
GLELO	*M. alpina*		GLA, STA	318	AAF70417
MAELO	*M. alpina*			317	AAF71789
IgASE1	*I. galbana*		LA, ALA	263	AF390174
PSE1	*P. patens*		GLA, STA	290	AF428243

Source of origin of these proteins is indicated, as well as the substrate specificity, protein length, and NCBI accession number.

[76] that these enzymes were proven to be involved in the synthesis of long-chain fatty acids. ELO2 is involved in the synthesis of saturated and monounsaturated fatty acids of up to 24 carbon atoms in length, while ELO3 elongated a broader group of saturated and monounsaturated fatty acids, and is essential for the conversion of C24:0 to C26:0 [76]. All three ELO gene products identified so far are condensing enzymes of the elongation system, but they have different substrate preferences. Dittrich et al. [32] have examined in details the substrate specificity, cofactor requirement and kinetic characteristics of yeast elongation enzymes.

Recently, progress has been made in identifying the other components of the elongation system. YBR159ω was identified as encoding a β-ketoacyl reductase, a protein that catalyzes the second step in each cycle of fatty acid elongation [7,40]. The deduced protein sequence of YBR159ω gene is 347 amino acids, and has significant homology to orthologues present in *Drosophila melanogaster*, *C. elegans*, and *Arabidopsis* sp. The encoded protein is also a member of the short-chain alcohol dehydrogenase superfamily and shows limited homology to human steroid dehydrogenases. Although YBR159ω is a non-essential gene, the oxidoreductase Ybr159p is specifically required for heterologous elongation activity in *S. cerevisiae*, and this activity can be enhanced by the over-expression of itself.

TSC13 gene, which is essential for yeast viability, has also been identified as encoding an enoyl reductase, a protein that catalyzes the last step in each cycle of fatty acid elongation [56]. The deduced protein sequence of TSC13 gene is 310 amino acids, is evolutionarily conserved, and has significant homology to rat, human, and *Arabidopsis* steroid reductases. The encoded protein, Tsc13p, co-immunoprecipitates with Elo2p and Elo3p, is localized to endoplasmic reticulum, and is enriched at sites of nuclear-vacuolar interface.

2.3. Elongation in mammals

The first identified mammalian gene encoding the elongation enzyme was thought to be involved in the recruitment of brown adipose tissue in mice exposed to cold. The enzyme was named as Cig30 (cold-inducible glycoprotein of 30 kDa) [115]. Based on homology to Cig30, Tvrdik et al. [116] have subsequently identified two other mouse genes, Ssc1 and Ssc2. All three enzymes were later reclassified as belonging to a gene family (ELOVL) involved in the biosynthesis of very long chain fatty acids. Complementation studies in yeast mutants indicated that Ssc1 (Elovl1) was functionally equivalent to yeast ELO3, and that Cig30 (Elovl3) was equivalent to ELO2. They are involved in the elongation of a broad group of saturated and monounsaturated fatty acids up to 24 carbons. The specific activity of Scc2 (Elovl2) was not characterized at that time. However, Tvrdik et al. [116] suggested that since the level of expression of Elovl2 is high in testis, which is also rich in long-chain PUFA, Elovl2p is likely involved in the metabolism of long-chain PUFA species. This view has recently been confirmed in our study in which we identified this mouse gene by expression in mammalian and yeast systems and showed its involvement in the elongation of both C_{20} and C_{22} long-chain PUFA [64]. A list of the various mammalian elongases along with their proposed activity is displayed in Table 1.

In mammalian cells, fatty acids with lengths of 16-18 carbon atoms, constitute a majority of cellular fatty acids. They are major products of de novo synthesis by the activity of cytoplasmic fatty acid synthase (FAS). Since elongation of fatty acids by FAS ends at palmitic acid (C16:0), and the end product of mammalian lipogenesis is usually oleic acid (C18:1n-9) or vaccenic acid (C18:1n-7), there is a missing step in the conversion of C_{16} to C_{18} fatty acids. Recently, Moon et al. [72] have identified and characterized a mouse cDNA corresponding to a mammalian fatty acid elongase by microarray analysis of mRNA obtained from the sterol regulatory element-binding protein (SREBP)-transgenic mouse livers, and designated it "long chain fatty acyl elongase" (Lce). Expression of the LCE gene in transgenic mice revealed that all tissues contained Lce mRNA, with high fat-containing tissues such as brown adipose tissue, white adipose tissue, liver, and brain, displaying high levels of Lce transcripts, whereas the low fat-containing tissues like spleen, skeletal muscle, and heart exhibiting low levels of the Lce transcript [72]. The same enzyme has also been identified and characterized by Matsuzaka et al. [67] using DNA microarray. They named the enzyme as fatty acyl-CoA elongase (FACE). In vitro microsomal fatty acid elongation assays demonstrated that FACE possesses elongase activity specific for C_{16} saturated and monounsaturated fatty acids (e.g., C16:0 and C16:1n-7). The enzyme is activated by the sterol regulatory element-binding protein (SREBP) with its expression being induced in the liver and adipose tissue in the refed state after fasting, but suppressed by the presence of dietary PUFAs [67]. Evidence has shown that mouse Lce (FACE) is a condensing enzyme in the elongation system. It encodes 267 amino acid residues and has 96% identity with the human homologue of 265 amino

acids [72]. It contains five transmembrane regions, a histidine-rich motif (HXXHH), and ER retention signal (KKXX-like), all typical for members of fatty acid elongase family [21,76,116,130].

The first genes encoding enzymes specific for the elongation of long-chain polyunsaturated fatty acids were identified in our laboratories from human and mouse [63]. The human HELO1 cDNA sequence was identified from a human EST library based on the limited identity with the *S. cerevisiae* ELO2 sequence. When the HELO1 cDNA was expressed in yeast cells, its protein product converted a wide range of exogenously added C_{18} and C_{20} long-chain PUFA substrate into their respective C_{20} and C_{22} elongated products: GLA into DGLA, AA into adrenic acid (ADA, C22:4n-6), STA into eicosatetraenoic acid (ETA, C20:4n-3), EPA into DPA and ALA into eicosatrienoic acid (ETrA, 20:3n-3). It is postulated that this enzyme is also involved in the elongation of ω-3 DPA to ω-3 TPA during the biosynthesis of DHA in mammals (Fig. 1) [105-107], although activity studies using C_{22} PUFAs as substrates were not carried out at the time due to un-availability of C_{22} PUFA substrate. Quantitative PCR demonstrated the distribution of the HELO1 transcript in 22 different human tissues. The highest levels of HELO1 mRNA were found in the testis and adrenal gland. Brain, lung, and prostate tissue also showed significant amounts of this transcript. These results are consistent with the fact that these tissues contain very high levels of ADA, an elongation product of AA [36,113]. The HELO1 is also a member of the ELOVL family. The HELO1 (ELOVL5) encodes 299 amino acids [63]. When the predicted amino acid sequence encoded by HELO1 was compared with those of other enzymes involved in the elongation of fatty acids, the human enzyme showed only 29.1% identity with yeast ELO2, and 27.7% identity with yeast ELO3.

Very long chain fatty acids (VLCFA) are fatty acids with greater than 20 carbon atoms. VLCFA can be divided into saturated, monunsaturated and polyunsaturated. Polyunsaturated VLCFA up to 36 carbons have been detected in mature spermatozoa [84], retina [3], and also in brain [83,95,98,112]. Sauerwald et al. [95] have shown the presence of radiolabeled 22:4n-3 and 24:4n-3 in the plasma phospholipid fraction of infants receiving a dose of [U–^{13}C]18:3n-3 and suggested that there exists an elongase enzyme capable of producing 22:4n-3 and 24:4n-3 through sequential addition of two carbon units to carboxyl end of polyunsaturated fatty acid precursors. In human, the gene responsible for elongation of VLCFA is ELOVL2 and ELOVL4 [64,130].

The human ELOVL2, as well as the mouse homologue Elovl2 were identified by blast analysis of NCBI database using the ELOVL5p sequences as the query. The ELOVL2 gene encodes a protein of 296 amino acids, sharing 56.4% amino acid sequence identity with ELOVL5 [64]. Expression studies in yeast revealed that this gene could efficiently elongate C_{20} and C_{22} PUFA substrates, but could not elongate C_{18} PUFA substrates, monounsaturated fatty acids or saturated fatty acids. The mouse homologue Elovl2 also showed specificity for C_{20} and C_{22} PUFA substrates when expressed in mouse L-cells. Northern analysis revealed that the highest level of expression of Elovl2 was found in mouse testis and liver, followed by brain and kidney [64].

ELOVL4 was identified by linkage and haplotype analysis in families with two forms of autosomal dominant macular dystrophy. Genetic mapping indicated that the ELOVL4 gene encodes a protein that is 44% identical to Ssc2 and 35% identical to members of ELO gene family in yeast. Human ELOVL4 encodes for 314 amino acids with 92% identity to the mouse orthologue. Northern blot analysis revealed that ELOVL4 transcript was abundant in the human retina, exclusively being expressed in the photoreceptor cells, and also in brain and testis. All of these

tissues are rich in very long chain fatty acids. Since the ELOVL4 has high homology to proteins that are involved in the elongation of long chain fatty acids, Zhang et al. [130] have suggested that the encoded enzyme is involved in the elongation of very long chain fatty acids.

A recent PCT publication by Winther et al. [123] described the identification, isolation, and expression of seven human elongase genes in yeast. All seven encoded enzymes were classified as PUFA elongases. The ELG2 (ELOVL1), which is 92% identical to the mouse Elovl1, has been shown to elongate γ-linolenic acid (GLA; C18:3n-6), AA, ALA, and EPA by two carbon chains to their respective products [123]. The ELG3 gene, an orthologue of ELOVL5 (HELO1) has been shown to be active when co-expressed with the human Δ5- and Δ6-desatuases [31]. The ELG4 gene elongates GLA, AA, ALA, and EPA. The ELG6 gene, a human orthologue of ELOVL3, has also been shown to have slight activity in elongating GLA and ALA when expressed in yeast [123].

In rat, two elongase genes have been identified, rELO1 and rELO2 [47]. The rELO1 gene is a homologue of human ELOVL5, which is also 299 amino acids, was shown to elongate mono- and polyunsaturated fatty acids of C_{16}–C_{20}. Northern analysis showed that rELO1 was expressed constitutively in various tissues with the highest levels observed in the brain and the lung [47]. Fasting and refeeding cycles of rats did not significantly affect the abundance of rELO1 in Northern analysis. The rELO1 sequence had high similarity with the sequence of ELOVL5 (93%), followed by mouse Elovl2 (57.4%) and Elovl3 (25.8%). The rELO2 gene, whose encoded protein contains 267 amino acids, was demonstrated to elongate C16:0 and to a lesser extent C18:0, and fatty acids with a low degree of desaturation. Northern analysis revealed that rELO2 expression in hepatocytes was activated by the fasting and refeeding cycles of rats on fat-free diet [47]. This contrasts with the constitutional expression of rELO1 suggesting that a different regulatory mechanism is responsible for the expression of the two elongases. The rELO2 gene has 95.8% identity with the human homologue, ELOVL3 and 42.9% identity with mouse Elovl3.

2.4. Elongation in plants

In plants, there also exist very-long-chain fatty acids (VLCFAs) with chain lengths of 20-carbon or more. In contrast to the polyunsaturated VLCFA seen in animals however, VLCFAs in plants are saturated or monounsaturated, and are used as constituents of surface coverings, such as waxes [82] or as seed storage lipids [59]. The basic chemistry of fatty acid elongation in plants is similar to that in animals. Fatty acid elongation is divided into four different reactions: condensation of malonly-CoA with a long-chain acyl-primer to form a β-ketoacyl-CoA, reduction to β-hydroxyacyl-CoA, dehydration to *trans*-2-enoyl-CoA and further reduction resulting in the elongated acyl-CoA [34]. The initial condensation step is catalyzed by the membrane-bound fatty acid elongase 1 (FAE1) also known as β-ketoacyl CoA synthase (KCS). The FAE1 gene which is involved in the synthesis of the 22-carbon monounsaturated fatty acid, erucic acid (C22:1n-9) was first identified in *Arabidopsis thaliana* [50,59,62]. The FAE1 homologues have also been found in other species such as *Brassica* sp. and jojoba [5,60]. The *Arabidopsis* FAE1 gene was subsquently cloned and characterized [28,51,71,118]. Millar and Kunst [71] showed that FAE1 encodes the condensing enzyme, β-ketoacyl-CoA synthase (KCS), which regulates the substrate specificity and activity of the elongation process. All plant FAE1 enzymes are specific for saturated and monounsaturated fatty acids, but not polyunsaturated fatty acids. None of the condensing

enzymes share a significant homology with the yeast ELO genes or the mammalian PUFA elongases. The KCS and ELO proteins differ from each other in a number of different aspects. The ELO proteins are very hydrophobic containing 5–7 transmembrane domains spread evenly between the N- and C-terminus (refer to Section 2.5.3). The KCS proteins on the other hand have a small membrane anchor of 1–2 transmembrane helices confined to the N-terminus. They also have a conserved active site cysteine-residue that is not seen in the ELO proteins [128].

2.5. Elongation in lower eukaryotes

2.5.1. Nematode

Caenorhabditis elgans is a free-living nematode. *C. elegans* can synthesize a wide range of PUFAs, including AA and EPA, by converting saturated fatty acids obtained from their diet [46]. This worm possesses the full range of desaturase activities (Δ12, ω15, Δ5, and Δ6-desaturases), in addition to their n-6 and n-3 PUFA elongase activities. Beaudoin et al. [8] characterized and expressed the *C. elegans* ORF (F56H11.4) encoding the presumptive condensing enzyme activity of a fatty acid elongase in yeast. Designated as CEELO1, the 288 amino acid elongase of ORF F56H11.4 was also identified by searching the database with the *M. alpina* elongase sequence [30]. This elongase shows specificity for the C_{18}-PUFAs GLA and SDA when expressed in yeast [8]. This ORF shows some limited homology to the yeast ELO gene family required for very long chain saturated fatty acid elongation [76] and some homology with mouse Cig30 (Elovl3) gene [115].

The heterologous reconstitution of the PUFA biosynthetic pathway in yeast was demonstrated by co-expression studies of the *C. elegans* elongase with a Δ5- or Δ6- desaturase [8]. The genomic and functional characterizations of PUFA biosynthesis in *C. elegans* have also recently been addressed [74]. These knock-out mutation studies demonstrated that *C. elegans* does not require n-3 or Δ5-unsaturated PUFAs for normal development. The persistence of n-3 and n-6 elongated C_{20} fatty acids in the elongase knock-out mutants, however, indicated that at least one additional elongase is present in this nematode [122]. Unlike human, *C. elegans* lacks the ability to elongate C_{20} PUFAs, resulting in the biosynthetic pathway ending with EPA, the most abundant PUFA in this invertebrate.

2.5.2. Fungus

Many fungi, particularly the oleagenous species, have the ability to synthesize very long-chain PUFAs. For example members of the *Saprolegnia* sp. are rich in AA and EPA, *Conidiobolus* sp. and *M. alpina* are rich in AA [55]. *M. alpina* can accumulate up to 40% (by wt.) lipid, of which up to 40% represents AA. Thus, this organism is commercially used in production of AA. Wynn and Ratledge [125] examined the fatty acid biosynthesis in *M. alpina* by studying its growth on lipid-based carbon sources. They showed that regulation of fatty acid synthesis and fatty acid desaturation/elongation was separate and not repressed by the growth of the fungus on simple fatty acids. Based on their data, they concluded that the elongation of C18:3n-6 to C20:3n-6 is the rate-limiting step in the overall conversion of oleic acid to AA in *M. alpina*. The elongase responsible for this conversion, designated as GLELO, wasidentified in our laboratories by expression studies of the *M. alpina* EST library in yeast, and following the conversion of C18:3n-6 to C20:3n-6 [29,30,77]. The encoded protein of 318 amino acids produced AA and EPA in yeast when co-expressed with the *M. alpina* Δ5-desaturase. Another *M. alpina* elongase, which has 40% identity

to yeast Elo2p, was also identified. This second elongase (MAELO) had similar activities as Elo2p in elongating saturated and monounsaturated fatty acids of 16-18 carbons and very weak activity on PUFA substrates [29,30,77].

2.5.3. Marine protist

Marine protests like the *Thraustochytrid* family of organisms, make large amounts of DHA, with some species accumulating >40% of the total lipids as DHA [55]. *Schizochytrium*, a member of the *Thraustochytrid* family, was recently demonstrated to utilize a polyketide synthase (PKS) pathway as its fatty acid biosynthesis system, similar to that employed by prokaryotes for PUFA biosynthesis [68]. In this case, the authors ruled out the existence of a functional membrane bound desaturase/elongase system for biosynthesis of DHA based on the low proportion of desaturase genes identified during the random sequencing of the EST library. In addition, whole cell lipid analysis revealed a lack of C_{18}-, C_{20}-, and C_{22}-PUFA intermediates, which is typically seen in organisms utilizing the PKS pathway for PUFA production [68]. This generalization, however, can not be applied to all members of the *Thraustochytrid* family with the recent discovery of a Δ4-desaturase from a *Thraustochytrium* sp. [87]. In addition, a Δ5-desaturase was also identified from this organism [87]. Recently, Heinz et al. [42] have identified an elongase-like sequence from a *Thraustochytrium* sp., and yeast expression studies have demonstrated that this enzyme can elongate C_{18} PUFA substrates. These findings implicate that these marine protists may have two pathways for DHA production, and thus may possess all the enzymes of the desaturase/elongase pathway required for synthesis of DHA, along with the PKS system for DHA production.

2.5.4. Microalga

Many microalgae are rich in long-chain n-3 PUFAs. For example, *Porphyridium cruentum*, *Nannochloropsis* sp., *Phaedactylum tricornutum* and *Monodus subterraneus* are rich in EPA, whereas *Crypthecodinium cohnii*, *Chroomonas salina* and *Isochroysis galbana* are rich in DHA [22,23,43,52,109,127]. The green oleaginous alga, *Parietochloris incisa*, is one of the few microalga that is rich in arachidonic acid [10].

A recent study has identified a C_{18} elongase in *I. galbana* by PCR-based strategy [86]. Designated as IgASE1, the encoded protein is predicted to be 263 amino acids and 30 kDa. The protein shares only limited homology to animal and fungal proteins with elongating activity. It is the first published elongase that substitutes a glutamine (Q) for the first histidine of the histidine box motif. Expression in yeast cells indicate that this protein elongates specifically linoleic acid and α-linolenic acid to eicosadienoic acid (C20:2n-6) and eicosatrienoic acid (C20:3n-3), respectively. Qi et al. [86] suggest that *I. galbana* may utilize the Δ8-desaturation pathway for its EPA and DHA synthesis.

2.5.5. Moss

The moss *Physcomitrella patens* contains high levels of fatty acids up to 20 carbons, and their biosynthesis involves the desaturation/elongation enzymes. *P. patens* elongase, PSE1, was characterized as a 290 amino acid protein. The protein sequence shares only 38.5% identity and 48.3% similarity with the yeast ELO1 [129], but has very strong identity with the *M. alpina* GLELO [128]. Expression studies of PSE1 in *S. cerevisiae* led to the elongation of GLA and STA,

and deletion mutants of PSE1 gene provided additional evidence for the function of PSE1. *P. patens* also has a β-ketoacyl-CoA synthases (KCS), similar to the KCS involved in the elongation system in plants [71]. In plants the KAS is shown to be specific for saturated and monounsaturated fatty acids, and thought to be involved in the supply of very long-chain fatty acids for waxes and seed storage lipids. Genome of other higher plants, such as *Arabidopsis*, corn, cotton, soybean, and wheat also contain both KCS-like and ELO-like sequences.

2.6. Sequence alignments/phylogenetic tree

The molecular identification of the fatty acid elongases involving in the biosynthesis of long-chain fatty acids has revealed their structural characteristics. There is a highly conserved motif, a histidine box containing three histidine residues, HXXHH, embedded in the fourth membrane spanning region, and the presence of five hydrophobic stretches predictable as a membrane-spanning region [63,67,72,116,123]. However, these features are not detected in plant enzymes, such as KAS II [119] and FAE1 [51]. Inagaki et al. [47] have speculated that the differences were

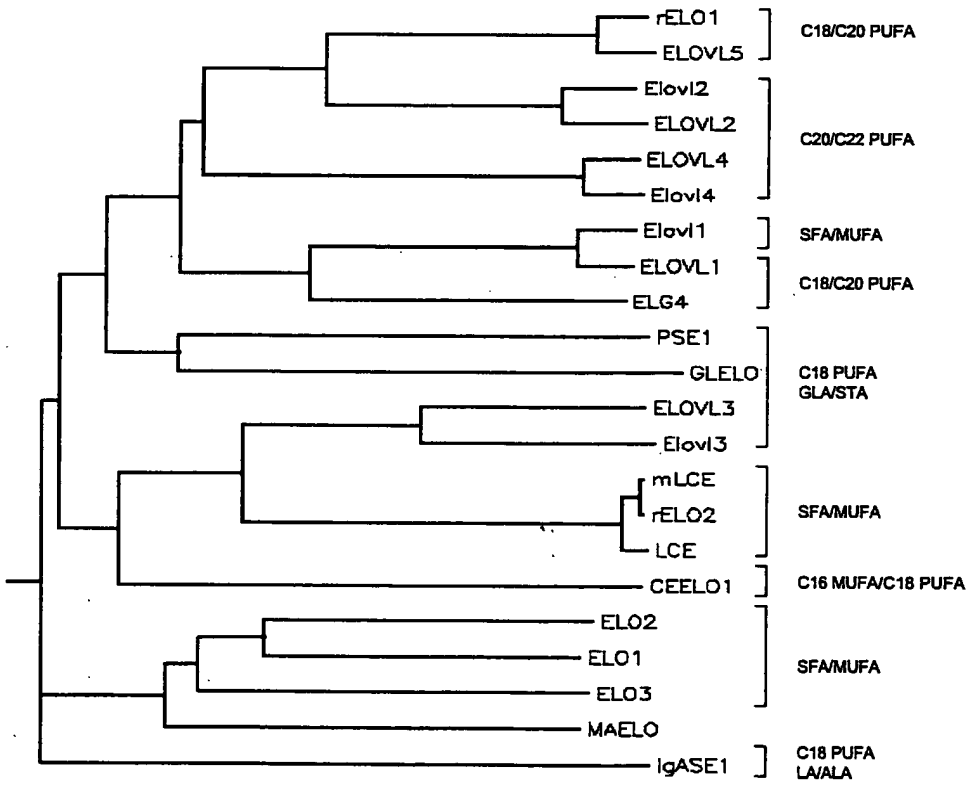


Fig. 3. A rooted phylogenic tree based on amino acid sequence identity of proteins belonging to the ELO family of elongating enzymes. The length of the branches is proportional to the evolutionary divergence among these proteins. The substrate specificity of these enzymes is indicated to the right. Proteins with common substrate preferences appear to group together.

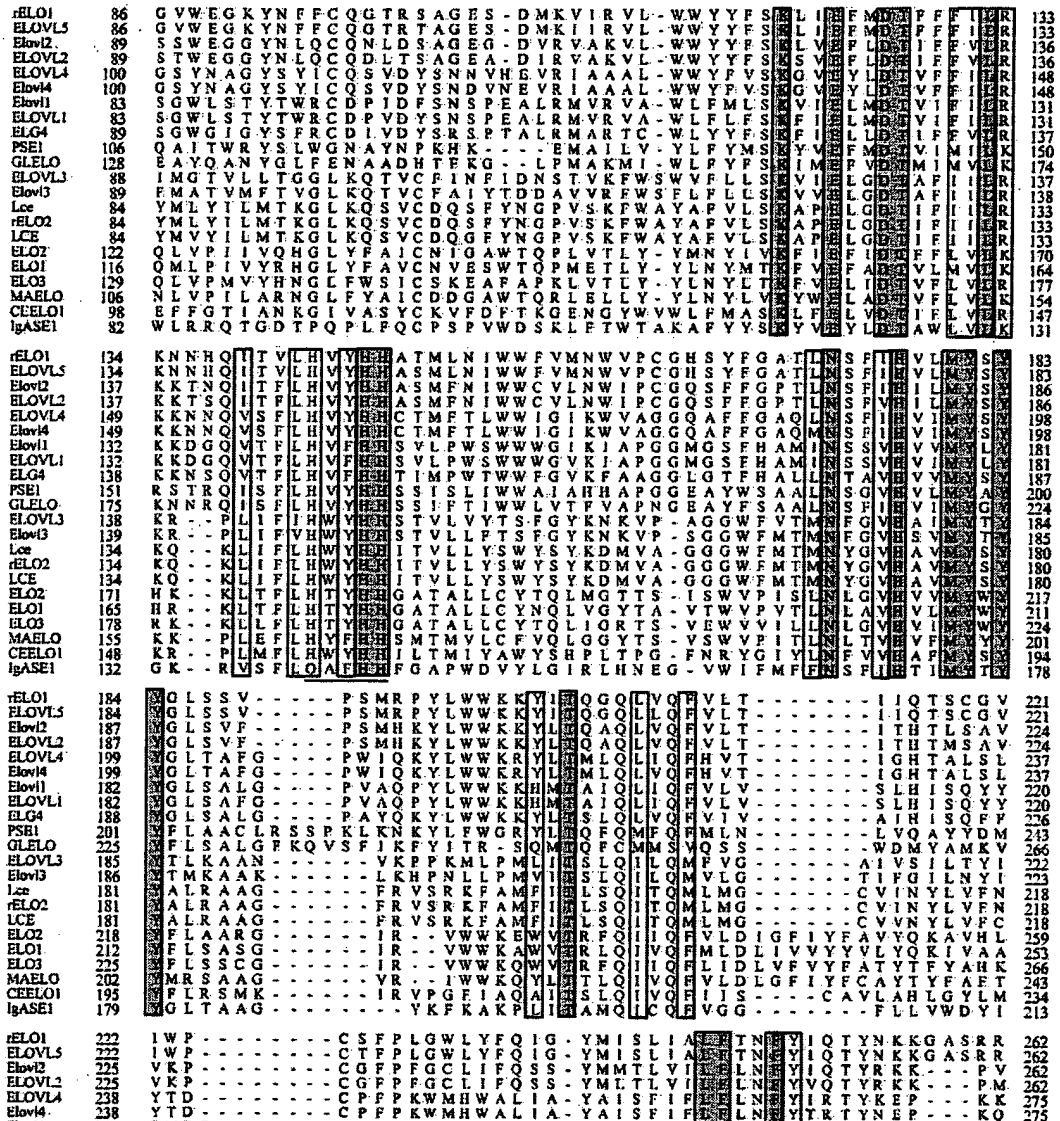


Fig. 4. ClustalW alignment of the deduced amino acid sequences of 22 members of the ELO family of elongating enzymes. These proteins are obtained from human (ELOVL, LCE), mouse (Elovl, Lce), rat (rELO), worm (CEELO), moss (PSE), fungi (GLELO, MAELO), and algae (IgASE). Only regions of the sequences that contain amino acid homology are depicted. Numbers indicate the position of the amino acid residues from the start of the sequence. Identical residues are shown in shaded boxes, and highly conserved residues in open boxes. The conserved Histidine box is underlined.

due to the differences in the structure of derivative groups. The plant enzymes use fatty acids bound to acyl carrier proteins (ACP) as a substrate [44], whereas elongases from animals (may include fungi, yeast), require the CoA-activated substrate.

The amino acid sequences of 90 enzymes expected or known to be involved in the elongation of fatty acids were aligned with ClustalX [128]. The protein diverged into two groups within the phylogenetic tree, with each group showing similarities to either ELO proteins or KCS proteins. Of the ELO proteins, 21 known elongase sequences were aligned with ClustalW. The phylogenetic tree shows the similarities of the mammalian homologues, with some divergence between the proteins of different unsaturation preferences (Fig. 3).

The ClustalW alignment of these 22 known elongases clearly highlights the conserved amino acids (Fig. 4). One highly conserved motif, a histidine box containing three histidine residues, found also in a number of membrane-bound desaturases, is common to all 22 sequences. Fifteen other fully conserved residues are identified in the multiple alignments where they are indicated by an asterisk (Fig. 4). These residues, as well as the conserved strong groups that are indicated by a semi-colon, suggest their critical importance in the catalytic activity and structure of the elongases. Winther et al. [123] proposed a model for the human elongases embedded in the endoplasmic reticulum (ER), the elongase encompasses a ring of transmembrane domains enclosing an inner catalytic cavity for insertion of fatty acyl chains.

3. Conclusion

Fatty acids are critical for the normal development and function of all organisms, and in particular, very long chain PUFAs are necessary for the health and maintenance of higher organism such as mammals. In the past, the extreme hydrophobicity of the identified elongase proteins has prevented their purification and thus the precise biochemical characterization of their roles in fatty acid elongation. However, much progress has been made in the study of the fatty acid biosynthesis pathways using molecular biology tools. Also the recent identification of a number of different elongases from various sources will provide a means for studying structure–function relationships based on sequence comparative analysis and mutation studies. Comparative sequence analysis has already helped group the condensing enzyme of the elongation system into two categories; the ELO system that includes enzymes involved in long chain PUFA production, and the KCS/FAE system found exclusively in plants for the production of waxes and seed storage lipids.

Although the long chain PUFA biosynthesis machinery was known to exist in various organisms, not much was known about the enzymes involved in this pathway, particularly the elongases. However with the advent of the genome sequencing projects for various organisms, many ELO like sequences have been identified from the database of various organisms including human, mouse, rat, *Arabidopsis*, *S. cerevisiae*, *L. major*, *D. melanogaster*, *M. fascicularis*, and *C. elegans*. Even for organisms whose genomes were not being sequenced, the presence of PUFAs, and the discovery of at least one enzyme from the desaturase/elongase pathway, provides grounds for the pursuit and identification of other enzymes that are involved in this biosynthesis process. For example, the discovery of a Δ6-desaturase in *Physcomitrella* confirmed the existence of the desaturase/elongase pathway, which led to the identification of the *PSE1*

encoding the elongase involved in PUFA production [37,128]. The ability to express the elongase genes in heterologous host expression systems have further contributed to our understanding of lipid function and modifications. In *S. cerevisiae* for example, the absence of an endogenous long chain PUFAs biosynthesis machinery provides for easy, uncomplicated screening system for identification of unknown genes with PUFA elongating activity. This system was used for the identification of the first elongase involved in PUFA biosynthesis [77]. Chuang et al. [20] have also used this system to demonstrate elongation of conjugated linoleic acid by HELO1p.

The identification of PUFA elongase from mammalian sources, as well as the other enzymes involved in the biosynthesis of long chain PUFAs, offers the potential of studying various fatty acid associated disorders and understanding their mode of action. Since diabetes and cardiovascular disease are all related to lipid metabolism dysfunction, interest in this area exploded over the past couple of years. In addition, inflammatory disease like arthritis, and also certain cancers, are shown to be associated with aberrant n-6/n-3 PUFA ratios, leading to overproduction of the prostaglandins and leukotrienes that enhance inflammation. Thus there is a drive to understand the regulation of the n-6/n-3 pathways, and identify compounds that can modulate the ratios of the different PUFAs.

Recent advances in the PUFA biosynthesis pathway has opened the avenue for better understanding of the various chronic diseases from the genetic stand point, as well as from the disease stand point. Genetic mapping data identified a 5-bp deletion in ELOVL4 as being associated with an inherited form of autosomal dominant macular dystrophy [130]. And based on recent identification and functional characterization of elongase genes from other organisms, the functionality of ELOVL4 was identified as an elongase.

The identification of different elonagase and desaturase genes involved in the EPA/DHA production has important biotechnological applications. These genes can be used in the production of PUFA-rich transgenic plant oils for therapeutic and prophylactic use. Also, advances in understanding gene regulation in PUFA biosynthesis will also impact the single-cell oil industry, such that growth conditions of the microalgae can be manipulated to enhance the production of EPA/DHA. This in turn will impact the marine fish-farming industry which depends on microalgae for enhancing the levels of PUFAs in fish. All the above will eventually afford the public an economical source of balanced n-3/n-6 PUFA-enriched oils that will greatly impact general health and nutrition in the near future.

References

[1] Abbadi A, Domergue F, Meyer A, Riedel K, Sperling P, Zank TK, et al. Eur J Lipid Sci Technol 2001;103:45–55.
[2] Aki T, Shimada Y, Inagaki K, Higashimoto H, Kawamoto S, Shigeta S, et al. Biochem Biophys Res Commun 1999;255:575–9.
[3] Aveldano MI, Sprecher H. J Biol Chem 1987;262:1180–6.
[4] Babcock T, Helton WS, Espat NJ. Nutrition 2000;16:1116–8.
[5] Barret P, Delourme R, Renard M, Domergue F, Lessire R, Delseny M, et al. Theor Appl Genet 1998;96:177–86.
[6] Bazan HE, Careaga MM, Sprecher H, Bazan NG. Biochim Biophys Acta 1982;712:123–8.
[7] Beaudoin F, Gable K, Sayanova O, Dunn T, Napier JA. J Biol Chem 2002;277:11481–8.

[8] Beaudoin F, Michaelson LV, Lewis MJ, Shewry PR, Sayanova O, Napier JA. Biochem Soc Trans 2000;28:661–3.
[9] Bernert Jr JT, Sprecher H. J Biol Chem 1977;252:6736–44.
[10] Bigogno C, Khozin-Goldberg I, Boussiba S, Vonshak A, Cohen Z. Phytochemistry 2002;60:497–503.
[11] Bourre JM, Parturneau JM, Duadu OL, Baumann N. Eur J Biochem 1977;72:41–7.
[12] Bourre JM, Dumont OS, Piciotti MJ, Pascal GA, Durand GA. Biochim Biophys Acta 1982;1124:119–22.
[13] Brenner RR. Prog Lipid Res 1982;20:41–7.
[14] Broun P, Boddupalli S, Somerville C. Plant J 1998;13:201–10.
[15] Carlson SE, Werkman SH, Peeples JM, Cooke RJ, Tolley EA. Proc Natl Acad Sci USA 1993;90:1073–7.
[16] Chapkin RS, Coble KJ. Biochim Biophys Acta 1991;1085:365–70.
[17] Charnock JS, Crozier GL, Woodhouse J. Nutr Res 1994;14:1089–99.
[18] Cho HP, Nakamura M, Clarke SD. J Biol Chem 1999;274:471–7.
[19] Cho HP, Nakamura M, Clarke SD. J Biol Chem 1999;274:37335–9.
[20] Chuang L-T, Leonard AE, Liu J-W, Mukerji P, Bray TM, Huang Y-S. Lipids 2001;36:139–43.
[21] Cinti DL, Cook L, Nagi MN, Suneja SK. Prog Lipid Res 1992;31:1–51.
[22] Cohen Z. J Am Oil Chem Soc 1990;67:916–20.
[23] Cohen Z. J Am Oil Chem Soc 1994;71:941–5.
[24] Connor WE. Am J Clin Nutr 2000;71:171 S-175S.
[25] Covello PS, Reed DW. Plant Physiol 1996;111:223–6.
[26] Crawford MA, Costeloe K, Ghebremeskel K, Phylactos A, Skirvin L, Stacey F. Am J Clin Nutr 1997;66:1032 S-1041S.
[27] Cunnane SC, Anderson MJ. J Nutr 1997;127:146–52.
[28] Das S, Roscoe TJ, Delseny M, Srivastave PS, Lakshmikumaran M. Plant Sci 2002;162:245–50.
[29] Das T, Thurmond JM, Bobik E, Leonard AE, Parker-Barnes JM, Huang Y-S, et al. Biochem Soc Trans 2000; 28:658–60.
[30] Das T, Thurmond JM, Leonard AE, Parker-Barnes JM, Bobik E, Chuang L-T et al. In: Huang YS, Ziboh VA, editors. γ-Linolenic acid—recent advances in biotechnology and clinical applications. Champaign, IL: AOCS Press; 2002. p. 40–4.
[31] De Antueno RJ, Knickle LC, Smith H, Elliot ML, Allen SJ, Nwaka S, et al. FEBS Lett 2001;509:77–80.
[32] Dittrich F, Zajonc D, Huhne K, Hoja U, Ekici A, Greiner E, et al. Eur J Biochem 1998;252:477–85.
[33] Falcone DL, Gibson S, Lemieux B, Somerville C. Plant Physiol 1994;106:1453–9.
[34] Fehling E, Mukerjee KD. Biochim Biophys Acta 1991;1082:239–46.
[35] Ferdinandusse S, Denis S, Mooijer PA, Zhang Z, Reddy JK, Spector AA, Wanders RJ. J Lipid Res 2001; 42:1987–95.
[36] Garcia PT, Holman RT. J Am Oil Chem Soc 1965;42:1137–41.
[37] Girke T, Schmidt H, Zähringer U, Reski R, Heinz E. Plant J 1998;15:39–48.
[38] Goldberg I, Schechter I, Bloch K. Science 1973;182:497–9.
[39] Griffiths G, Jackson F, Brechany EY, Christie WW, Stymne S, Stobart AK. Phytochemistry 1996;203:381–6.
[40] Han G, Gable K, Kohlwein SD, Beaudoin F, Napier JA, Dunn TM. J Biol Chem 2002;277:35440–9.
[41] Hastings N, Agaba M, Tocher DR, Leaver MJ, Dick JR, Sargent JR, et al. Proc Natl Acad Sci USA 2001; 98:14304–9.
[42] Heinz E, Zank T, Zaehringer U, Lerchl J, Renz A. Patent: WO 0159128-A, 16 August 2001.
[43] Henderson RJ, Sargent JR. Phytochemistry 1989;28:1355–62.
[44] Hlousek-Radojcic A, Imai H, Jaworski JG. Plant J 1995;8:803–9.
[45] Huang Y-S, Chaudhary S, Thurmond JM, Bobik Jr EG, Yuan L, Chan GM, et al. Lipids 1999;34:649–59.
[46] Hutzell PA, Krusber LR. Comp Biochem Physiol 1982;73B:517–20.
[47] Inagaki K, Aki T, Fukuda Y, Kawamoto S, Shiegeta S, Ono K, et al. Biosci Biotechnol Biochem 2002;66:613–21.
[48] Innis SM, Sprecher H, Hachey D, Edmond J, Anderson RE. Lipids 1999;34:139–49.
[49] Jamal GA, Carmichael HA, Weir AI. Lancet 1986;i:1098.
[50] James DW, Dooner HK. Theor Appl Genet 1990;80:241–5.

[51] James Jr DW, Lim E, Keller J, Plooy I, Ralston E, Dooner HK. Plant Cell 1995;7:309–19.
[52] Jiang GY, Chen F, Liang SZ. Process Biochemistry 1999;34:633–7.
[53] Jump D. J Biol Chem 2002;277:8755–8.
[54] Kang JX, Leaf A. Circulation 1996;94:1774–80.
[55] Kendrick A, Ratledge C. Lipids 1992;27:15–20.
[56] Kohlwein SD, Eder S, Oh C-S, Martin CE, Gable K, Bacikova D, et al. Mol Cell Biol 2001;21:109–25.
[57] Knutzon DS, Thurmond JM, Huang Y-S, Chaudhary S, Bobik Jr EG, Chan GM, et al. J Biol Chem 1998; 273:29360–6.
[58] Kremer JM, Lawrence DA, Jubiz W, DiGiocomo R, Rynes R, Bartholomew LE, et al. Arthritis Rheum 1990; 33:810–20.
[59] Kunst L, Taylor DC, Underhill EW. Plant Physiol Biochem 1992;30:425–34.
[60] Lassner MW, Lardizabal K, Metz JG. Plant Cell 1996;8:281–92.
[61] Lauritzen L, Hansen HS, Jorgensen MH, Michaelsen KF. Prog Lipid Res 2001;40:1–94.
[62] Lemieux B, Miguel M, Somerville C, Browns J. Theor Appl Genet 1990;80:234–40.
[63] Leonard AE, Bobik EG, Dorado J, Kroeger PE, Chuang L-T, Thurmond JM, et al. Biochem J 2000;350:765–70.
[64] Leonard AE, Kelder B, Bobik EG, Chuang L-T, Lewis CJ, Kopchick JJ, et al. Lipids 2002;37:733–40.
[65] Leonard AE, Kelder B, Bobik EG, Chuang L-T, Parker-Barnes JM, Thurmond JM, et al. Biochem J 2000; 347:719–24.
[66] Luthria DL, Sprecher H. Biochim Biophys Acta 1997;1345:221–30.
[67] Matsuzaka T, Shimano H, Yahagi N, Yoshikawa T, Amemiya-Kudo M, Hasty AH, et al. J Lipids Res 2002; 43:911–20.
[68] Metz JG, Roessler P, Facciotti D, Levering C, Dittrich F, Lassner M, et al. Science 2001;293:290–3.
[69] Michaelson LV, Lauzarus CM, Griffiths G, Napier JA, Stobart AK. J Biol Chem 1998;273:19055–9.
[70] Michaelson LV, Napier JA, Lewis M, Griffiths G, Lauzarus CM, Stobart AK. FEBS Lett 1998;439:215–8.
[71] Millar AA, Kunst L. Plant J 1997;12:121–31.
[72] Moon Y-A, Shah NA, Mohapatra S, Warrington JA, Horton JD. J Biol Chem 2001;276:45358–66.
[73] Napier JA, Hey SJ, Lacey DJ, Shewry PR. Biochem J 1998;330:611–4.
[74] Napier JA, Michaelson LV. Lipids 2001;36:761–6.
[75] Nugteren DH. Biochim Biophys Acta 1965;106:280–90.
[76] Oh C-S, Toke DA, Mandala S, Martin CE. J Biol Chem 1997;272:17376–84.
[77] Parker-Barnes JM, Das T, Bobik E, Leonard AE, Thurmond JM, Chuang L-T, et al. Proc Natl Acad Sci USA 2000;97:8284–9.
[78] Passorn S, Laoteng K, Rachadawong S, Tanticharoen M, Cheevadhanarak S. Biochem Biophys Res Commun 1999;263:47–51.
[79] Peyou-Ndi MM, Watts JL, Browse J. Arch Biochem Biophys 2000;376:399–408.
[80] Pirtle IL, Kongcharoensuntorn W, Nampaisansuk M, Knesek JE, Chapman KD, Pirtle RM. Biochim Biophys Acta 2001;1522:122–9.
[81] Pollet S, Bourre JM, Daudu O, Baumann N. CR Acad Sci Hebd Seances Acad Sci D 1971;273:1426–9.
[82] Post-Beittenmiller D. Annu Rev Plant Physiol Plant Mol Biol 1996;47:405–30.
[83] Poulos A. Lipids 1995;30:1–14.
[84] Poulos A, Sharp P, Johnson D, White I, Fellenberg A. Biochem J 1986;240:891–5.
[85] Prasad MR, Nagi MN, Ghesquier D, Cook L, Cinti DL. J Biol Chem 1986;261:8213–7.
[86] Qi B, Beaudoin F, Fraser T, Stobart AK, Napier JA, Lazarus CM. FEBS Lett 2002;510:159–65.
[87] Qiu X, Hong H, Mackenzie SL. J Biol Chem 2001;276:31561–6.
[88] Reddy AS, Nuccio ML, Gross LM, Thomas TL. Plant Mol Biol 1993;27:293–300.
[89] Retterstol K, Haugen TB, Christophersen BO. Biochim Biophys Acta 2000;1483:119–31.
[90] Riemersma RA, Wood DA, Butler S, Elton RA, Oliver M, Salo M, et al. Br Med J 1986;292:1423–7.
[91] Sakamoto T, Los DA, Higashi S, Wada H, Nishida I, Ohmori M, et al. Plant Mol Biol 1994;26:249–63.
[92] Sakuradani E, Kobayashi M, Ashikari T, Shimizu S. Eur J Biochem 1999;261:812–20.
[93] Sato N, Fujiwara S, Kawaguchi A, Tsuzuki M. J Biochem (Tokyo) 1997;122:1224–32.

[94] Sato M, Adan Y, Shibata K, Shoji Y, Sato H, Imaizumi K. World Rev Nutr Diet 2001;88:196–9.
[95] Sauerwald TU, Hachey DL, Jensen CL, Chen H, Anderson RE, Heird WC. Pediatr Res 1997;41:183–7.
[96] Sayanova O, Smith MA, Lapinskas P, Stobart AK, Dobson G, Christie WW, et al. Proc Natl Acad Sci USA 1997;94:4211–6.
[97] Schneiter R, Tatzer V, Gogg G, Leitner E, Kohlwein SD. J Bacteriol 2000;182:3655–60.
[98] Sharp P, Johnson D, Poulos A. J Neurochem 1991;56:30–7.
[99] Simopoulos AP. Am J Clin Nutr 1999;70:560 S-569S.
[100] Simopoulos AP. J Am Coll Nutr 2002;21:495–505.
[101] Singer P, Jaeger W, Voigt S, Thiel H. Prostaglandins, Leukotrienes Med 1984;15:159–65.
[102] Spector AA. Lipids 1999;34:1 S-3S.
[103] Sprecher H. Biochim Biophys Acta 1974;360:113–23.
[104] Sprecher H. Prog Lipid Res 1982;20:13–22.
[105] Sprecher H. Nutrition 12:S5-S7 1996.
[106] Sprecher H. Biochim Biophys Acta 2000;1486:219–31.
[107] Sprecher H, Luthria DL, Mohammed BS, Baykousheva SP. J Lipid Res 1995;36:2471–7.
[108] Spychalla JP, Kinney AJ, Browse J. Proc Natl Acad Sci USA 1997;94:1142–7.
[109] Sukenik A, Carmeli Y. J Phycol 1989;25:686–92.
[110] Suneja SK, Nagi MN, Cook L, Cinti DL. J Neurochem 1990;57:140–6.
[111] Suneja SK, Osei P, Cook L, Nagi MN, Cinti DL. Biochim Biophys Acta 1990;1042:81–5.
[112] Svennerholm L, Stallberg-Stenhagen S. J Lipid Res 1968;9:215–25.
[113] Takayasu K, Okuda K, Yoshikawa I. Lipids 1970;5:743–50.
[114] Toke DA, Martin CE. J Biol Chem 1996;271:18413–22.
[115] Tvrdik P, Asadi A, Kozak LP, Nedergaard J, Cannon B, Jacobsson A. J Biol Chem 1997;272:31738–46.
[116] Tvrdik P, Westerberg R, Silve S, Asadi A, Jakobsson A, Cannon B, et al. J Cell Biol 2000;149:707–17.
[117] Uauy R, Peirano P, Hoffman D, Mena P, Birch D, Birch E. Lipids 1996;31(Suppl.):167 S-176S.
[118] Venkateswari J, Kanrar S, Kirti PB, Malathi VG, Chopra VL. J Plant Biochem Biotech 1999;8:53–5.
[119] von Wettstein-Knowles P, Olsen JG, Arnvig K, Larsen S. In: Harwood JL, Quinn PJ, editors. Recent Advances in the biochemistry of plant lipids. Portland Press; 2001. p. 601–7.
[120] Wallis JG, Browse J. Arch Biochem Biophys 1999;365:307–16.
[121] Watts JL, Browse J. Arch Biochem Biophys 1999;362:175–82.
[122] Watts JL, Browse J. Proc Natl Acad Sci USA 2002;99:5854–9.
[123] Winther MD, Knickle LC, Martin H, de Antueno RJ, Jenkins DK, Nwaka SO, et al. World Intellectual Property Organization WO 02/44320A2, 2002.
[124] Woodman RJ, Mori TA, Burke V, Puddey IB, Wats GF, Beilin LJ. Am J Clin Nutr 2002;76:1007–15.
[125] Wynn JP, Ratledge C. Microbiol 2000;146:2325–31.
[126] Yadav NS, Wierzbicki A, Aegerter M, Caster CS, Perez-Grau L, Kinney AJ, et al. Plant Physiol 1993;103:467–76.
[127] Yongmanitchai W, Ward OP. J Am Oil Chem Soc 1992;69:584–90.
[128] Zank TK, Zahringer U, Beckmann C, Pohnert G, Boland W, Holtorf H, et al. Plant J 2002;31:255–68.
[129] Zank TK, Zahringer U, Lerchl J, Heinz E. Biochem Soc Trans 2000;28:654–7.
[130] Zhang K, Kniazeva M, Han M, Li W, Yu Z, Yang Z, et al. Nature Genetics 2001;27:89–93.
[131] Zurier R, Rossetti RG, Jacobson EW, DeMarco DM, Liu NY, Temming JE, et al. Arthritis Rheum 1996; 39:1808–17.

The variant 'his-box' of the C18-Δ9-PUFA-specific elongase IgASE1 from *Isochrysis galbana* is essential for optimum enzyme activity

Baoxiu Qi*, Thomas C.M. Fraser, Claire L. Bleakley, Elisabeth M. Shaw, A. Keith Stobart, Colin M. Lazarus

School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK

Received 15 April 2003; accepted 19 May 2003

First published online 18 June 2003

Edited by Richard Cogdell

Abstract **IgASE1, a C18-Δ9-polyunsaturated fatty acid-specific fatty acid elongase component from *Isochrysis galbana*, contains a variant histidine box (his-box) with glutamine replacing the first histidine of the conserved histidine-rich motif present in all other known equivalent proteins. The importance of glutamine and other variant amino acid residues in the his-box of IgASE1 was determined by site-directed mutagenesis. Results showed that all the variation in amino acid sequence between this motif in IgASE1 and the consensus sequences of other elongase components was required for optimum enzyme activity. The substrate specificity was shown to be unaffected by these changes suggesting that components of the his-box are not directly responsible for substrate specificity.**

Key words: Fatty acid elongating activity; Site-directed mutagenesis; Specificity; *Isochrysis galbana*

1. Introduction

Much interest is currently being focused on the very long chain polyunsaturated fatty acids (PUFAs) arachidonic acid (AA), eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) because of their involvement in early human retinal and brain development and disease prevention [1–5]. Their role as precursors of eicosanoids, a family of biological effectors involved in inflammatory responses, regulation of blood pressure and blood clotting, has also been established [6]. They are obtained either directly from the diet or synthesised from the ω6 and ω3 essential fatty acids linoleic acid (LA, C18:2n-6, Δ9,12) and α-linolenic acid (ALA, C18:3n-3, Δ9,12,15), respectively, by alternating desaturation and elongation reactions. Fatty acid elongation is a four-step process involving condensation, reduction, dehydration and a second reduction reaction, with substrate specificity residing in the rate-limiting condensation reaction [7]. This means that elongases with different substrate specificities can be assembled using variable condensing components and common dehydration and reduction components. Studies on the mouse *LCE* gene indicated that it encodes the condensation component of a long-chain fatty acid elongase [8], and related genes have been isolated from a number of sources including yeast [9,10], *Mortierella alpina* [11] human [12], nematodes [13] and moss [14]. The characterisation of a C18-Δ9-specific PUFA elongase condensation component from *Isochrysis galbana* [15], together with the isolation of Δ8 and Δ4 desaturases from *Euglena* [16] and *Thraustochytrium* [17], suggested that EPA and DHA may be synthesised from ALA by a so-called ω3-Δ8 pathway in these organisms [15]. In this pathway, which is shown in Fig. 1, ALA is first elongated by a C18-Δ9-specific fatty acid elongase to C20:3n-3 (ω3-eicosatrienoic acid (EtrA)), with further desaturations/elongations yielding C20:5n-3 (EPA) and finally C22:6n-3 (DHA).

Alignment of the amino acid sequences of fatty acid elongase condensation components reveals the presence of a highly conserved histidine-rich motif, HXXHH ('his-box') (Fig. 2). In IgASE1, however, glutamine (Q) replaces the first histidine in the his-box, and the significance of this substitution is unclear [15]. Similarly, alignment of most membrane-bound fatty acid desaturases reveals three highly conserved his-boxes, and these have been implicated in the binding of di-iron, a requirement for catalytic activity [18]. Site-directed mutagenesis of these histidine residues resulted in enzyme inactivation [19,20]. In the so-called 'front-end' desaturases from plants, animals and fungi, the first histidine in the third his-box is also substituted with glutamine. Site-directed mutagenesis of this glutamine in the Δ6-fatty acid desaturase from borage resulted in a complete loss of enzyme activity even when it was replaced by histidine [21].

We have now assessed the importance of the conserved histidine box present in the elongase component, particularly the H to Q substitution found in IgASE1 of *I. galbana*. Using a polymerase chain reaction (PCR)-based site-directed mutagenesis strategy the glutamine in the histidine box of IgASE1 was converted to histidine, alanine or phenylalanine. Further changes generated the HVYHH sequence found in GLELO1 from *Mortierella alpina*, which is specific for elongating the two Δ6-desaturated PUFAs, γ-linolenic acid (GLA, C18:3n-6) and stearidonic acid (STA, C18:4n-3) [11].

2. Materials and methods

2.1. Mutagenesis by PCR

The amplification of the IgASE1 coding region from an *I. galbana*

*Corresponding author. Fax: (44)-117-9257374.
E-mail address: baoxiu.qi@bristol.ac.uk (B. Qi).

Abbreviations: PUFA, polyunsaturated fatty acid; AA, arachidonic acid; EPA, eicosapentaenoic acid; DHA, docosahexaenoic acid; LA, linoleic acid; ALA, α-linolenic acid; GLA, γ-linolenic acid; STA, stearidonic acid; EDA, ω6-eicosadienoic acid; EtrA, ω3-eicosatrienoic acid

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00676-8

ω3 (Δ8) pathway:

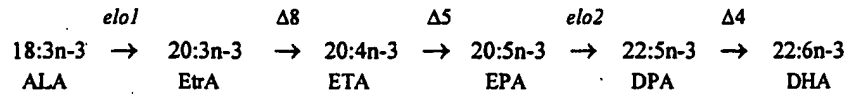


Fig. 1. Biosynthesis of EPA and DHA via the ω3 (Δ8) desaturation pathway by some microalgae. In this pathway, ALA is first elongated by a C18-Δ9-specific fatty acid elongase (elo1) to C20:3n-3 (EtrA). A Δ8 desaturase is required to add a double bond at the Δ8 position of the carbon chain to generate C20:4n-3 (ETA). Further desaturation by a Δ5 desaturase results in EPA (20:5n-3). DHA (C22:6n-3) is the product of one more elongation step (elo2) plus a Δ4 desaturation step.

cDNA library and its cloning into plasmid pCR2.1-TOPO (Invitrogen) is described in Qi et al. [15]. To change the his-box from QAFHH to HAFHH, first round PCRs were carried out with *Pfu* polymerase (Promega) using primer pairs M13 reverse (universal primer located on the vector) with GlnHisFor (5′-C TCC TTT CTC CAT GCC TTC CAC CAC-3′, and M13 forward with GlnHisRev (5′-GTG GTG GAA GGC ATG GAG AAA GGA G-3′). A second round PCR, containing 1 μl of each of the gel-purified first-round products, M13 forward and reverse primers and the Expand High Fidelity enzyme (Roche), was used to assemble the mutated IgASE1 coding region. The same techniques were used to change glutamine to alanine (QAFHH→AAFHH) or phenylalanine (QAFHH→FAFHH), alanine/phenylalanine to valine/tyrosine (QAFHH→QVYHH) and glutamine/alanine/phenylalanine to histidine/valine/tyrosine (QAFHH→HVYHH). The primers used are listed in Table 1. All PCR reactions were carried out under the following conditions: one initial denaturation step of 94°C for 3 min; 10 cycles of 94°C for 15 s, 55°C for 30 s and 72°C for 90 s; 20 cycles of 94°C for 15 s, 55°C for 30 s, and 72°C for 93 s (increasing by 3 s in each successive cycle); one final extension step of 72°C for 6 min.

Second round PCR products were cloned using the TOPO TA cloning system (Invitrogen) and the desired base changes confirmed by sequencing. Modified IgASE1 coding regions were excised from the pCR2.1-TOPO vector using *Kpn*I and *Eco*RI and ligated into the corresponding restriction sites of the yeast expression vector pYES2 (Invitrogen), downstream of the *GAL1* promoter. Subsequent yeast transformation and feeding experiments were performed according to Qi et al. [15].

2.2. Fatty acid analysis

Yeast cells were pelleted, washed and dried under a stream of N_2. Total fatty acids were extracted and transmethylated with methanolic HCl. The fatty acid methyl esters were analysed by gas chromatography (GC) on a 30 m×0.25 mm fused silica DB-23 capillary column (J&W Scientific) using heptadecanoic acid as internal standard and quantified by flame ionisation detection (FID). The chromatograph was programmed for an initial temperature of 140°C for 5 min followed by a 20°C/min temperature ramp to 185°C and a secondary ramp of 1.5°C/min to 220°C. The final temperature was maintained for 2 min. Injector and detector temperatures were maintained at 230°C and 250°C respectively. The initial head pressure of the carrier gas (He) was 90 kPa; a split injection was used.

3. Results and discussion

3.1. Functional analysis of wild-type IgASE1

Plasmid pY2ASE1 contains the IgASE1 open reading frame (ORF) under the control of the *GAL1* promoter in the yeast expression vector pYES2. Transformed yeast cells harbouring this plasmid were grown on minimal medium supplemented with LA and ALA to compensate for the lack of endogenous substrates for the C18-Δ9-specific elongase. After 48 h the fatty acid elongation products ω6-eicosadienoic acid (EDA, 20:2n-6) and EtrA (20:3n-3) accumulated to 13.6 and 14.8 mol% of total fatty acids, representing some 55 and 48% conversion of C18:2n-6 and C18:3n-3 to C20:2n-6 and C20:3n-3, respectively (Table 2). These data are consistent with our previous assays [15], where we also showed no elongase activity with GLA (20:3n-6) and clearly demonstrate, therefore, that the *IgASE1* gene encodes a C18-Δ9-specific PUFA elongating activity.

3.2. Mutagenesis of the IgASE1 his-box

Glutamine in the IgASE1 his-box was replaced by histidine, alanine or phenylalanine, and the mutant proteins were assayed for activity in transformed yeast. Table 3 shows that all the substitutions resulted in lower elongase activity, although complete inactivation was never achieved. The greatest activity was obtained with the alanine substitution, where more than 70% of the original enzyme activity was retained. The phenylalanine substitution had the lowest (but still measurable) activity, whilst the histidine substitution resulted in an activity that was some 50% of the control value. The glutamine residue in the histidine box thus appears to be essential for optimum enzyme catalysis, although its substitution did not result in complete enzyme inactivation. This is in contrast to the effect of similar mutagenic changes to the glutamine residue in the third his-box of a Δ6 desaturase from borage, which resulted in complete loss of enzyme activity [21]. Stepwise mutagenesis was also performed to change residues in the IgASE1 his-box (QAFHH) to match those found in the consensus his-box sequence of several other PUFA-specific elongases, including that of the GLA-specific elongase GLELO1 from *Mortierella* (HVYHH) (Fig. 2). The double mutant, in which the middle two amino acids alanine and phenylalanine in the histidine box were converted to valine and tyrosine but with the glutamine remaining unchanged (QAFHH→QVYHH), reduced the activity to less than 30% of the wild-type activity. The triple mutant, in which the first three amino acids in the histidine box, QAF, were replaced by HVY (QAFHH→HVYHH), had less than 10% of the wild-type

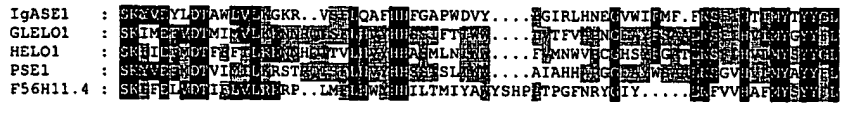


Fig. 2. Alignment of 'his-box' regions of the predicted protein sequence of *I. galbana IgASE1* [15] with sequences predicted from other known PUFA-specific elongating activity genes. *GLELO1* is from *Mortierella alpina* [11]; *HELO1* [12] is from human; *PSE1* is from *Physcomitrella patens* [14]; *F56H11.4* is from *Caenorhabditis elegans* [13].

Table 1
Oligonucleotide primers used in PCR reactions to construct site-directed mutants of IgASE1/pCR2.1-TOPO

Mutation	Oligonucleotide and amino acid sequences	His-box
Q→H	5'-C TCC TTT CTC CAT GCC TTC CAC CAC-3' S F L *H* A F H H	HAFHH
Q→A	5'- C TCC TTT CTC GCG GCC TTC CAC CAC-3' S F L *A* A F H H	AAFHH
Q→F	5'-GG GTC TCC TTT CTC TTC GCC TTC CAC CAC TTT G-3' V S F L *F* A F H H F	FAFHH
AF→VY	5'-TTT CTC CAG GTC TAC CAC CAC TTT G-3' F L Q *V* *Y* H H F	QVYHH
QAF→HVY	5'-TTT CTC CAT GTC TAC CAC CAC TTT G-3' F L *H* *V* *Y* H H F	HVYHH

The changed nucleotide bases are indicated by bold type, and the corresponding amino acids are indicated by bold, italic type and underlined. The mutated ORFs were confirmed by sequencing. The plasmids were cloned into pYES2 and the resulted constructs were used for yeast transformation and feeding experiments.

activity. These results suggest that interactions between the glutamine and its adjacent amino acids are important for maximising enzyme activity. The various his-box-mutated constructs were also expressed in the presence of numerous other exogenous fatty acid substrates including GLA, STA and EPA, but no elongation products were observed (data not presented).

The results suggest that the his-box is important for catalysis and that all the differences between the his-boxes of other elongases and IgASE1 are required for optimum enzyme activity. However, the elements required to regulate substrate specificity remain to be identified.

References

[1] Carlson, S.E., Werkman, S.H., Peeples, J.M., Cooke, R.J. and Tolley, E.A. (1993) Proc. Natl. Acad. Sci. USA 90, 1073–1077.
[2] Gill, I. and Valivety, R. (1997) Trends Biotechnol. 15, 401–409.
[3] Crowford, M. (2000) Am. J. Clin. Nutr. 71, S275–S284.
[4] Lauritzen, L., Hansen, H.S., Jurgensen, M.H. and Michaelsen, K.F. (2001) Prog. Lipid Res. 40, 1–94.
[5] Thies, F., Garry, J.M.C., Yaqoob, P., Rerkasem, K., Williams, J., Shearman, C.P., Gallagher, P.J., Calder, P.C. and Grimble, R.F. (2003) Lancet 361, 477–485.
[6] Kinsella, J.E., Lokesh, B., Broughton, S. and Whelan, J. (1990) Nutrition 6, 24–44.
[7] Cinti, D.L., Cook, L., Nagi, M.H. and Suneja, S.K. (1992) Prog. Lipid Res. 31, 1–51.
[8] Moon, Y.A., Shah, N.A., Mohapatra, S., Warrington, J.A. and Horton, J.D. (2001) J. Biol. Chem. 276, 45358–45366.
[9] Toke, D.A. and Martin, C.E. (1996) J. Biol. Chem. 271, 18413–18422.
[10] Oh, C., Toke, A.D., Mandala, S. and Martin, C.E. (1997) J. Biol. Chem. 272, 17376–17384.
[11] Parker-Barnes, J.M., Das, T., Bobik, E., Leonard, A.E., Thurmond, J.M., Chuang, L., Huang, Y.S. and Mukerji, P. (2000) Proc. Natl. Acad. Sci. USA 97, 8284–8289.
[12] Leonard, A.E., Bobik, E.G., Dorado, J., Kroeger, P.E., Chuang, L.T., Thurmond, J.M., Parker-Barnes, J.M., Das, T., Huang, Y.S. and Mukerji, P. (2000) Biochem. J. 350, 765–770.
[13] Beaudoin, F., Michaelson, L.V., Hey, S.J., Lewis, M.J., Shewry, P.R., Sayanova, O. and Napier, J.A. (2000) Proc. Natl. Acad. Sci. USA 97, 6421–6426.
[14] Zank, T.K., Zähringer, U., Beckmann, C., Pohnert, G., Boland, W., Holtorf, H., Reski, R., Lerchl, J. and Heinz, E. (2002) Plant J. 31, 255–268.
[15] Qi, B., Beaudoin, F., Fraser, T., Stobart, A.K., Napier, J.A. and Lazarus, C.M. (2002) FEBS Lett. 510, 159–165.
[16] Wallis, J.G. and Browse, J. (1999) Arch. Biochem. Biophys. 365, 307–316.
[17] Qiu, X., Hong, H. and MacKenzie, S.L. (2001) J. Biol. Chem. 276, 31561–31566.
[18] Shanklin, J. and Cahoon, E.B. (1998) Annu. Rev. Plant Physiol. Plant Mol. Biol. 49, 611–641.
[19] Shanklin, J., Whittle, E. and Cox, D.C. (1994) Biochemistry 33, 12787–12794.
[20] Avelange-Macherel, M.H., Tomita, T., Macherel, D., Wada, H. and Murata, N. (1995) FEBS Lett. 361, 111–114.
[21] Sayanova, O.V., Beaudoin, F., Libisch, B., Castel, A., Shewry, P.R. and Napier, J.A. (2001) J. Exp. Bot. 52, 1581–1585.

Table 2
Elongation products of different fatty acid substrates supplied to yeast cells transformed with the empty vector pYES2 and pY2ASE1

Fatty acid	Mol% of total fatty acids					
	pYES2			pY2ASE1		
	−substrate	+LA (18:2n-6)	+ALA (18:3n-3)	−substrate	+LA (18:2n-6)	+ALA (18:3n-3)
16:0	27.0	27.0	24.7	24.5	23.2	22.2
16:1n-9	41.7	35.1	30.1	43.4	32.3	24.1
18:0	6.0	5.7	5.0	5.8	5.3	5.0
18:1n-9	25.3	22.3	18.4	26.2	14.6	18.1
18:2n-6*	–	10.0	–	–	11.0	–
18:3n-3*	–	–	21.8	–	–	16.3
20:2n-6	–	–	–	–	13.6	–
20:3n-3	–	–	–	–	–	14.8
Elongation (%)	0	0	0	0	55.3	47.6

Exogenous fatty acids supplied as substrates for elongation are indicated by an asterisk. The values given are expressed as mol% of total fatty acid methyl esters identified by GC and FID. In the case of elongated substrates, this is also expressed as a per cent elongation (product/product+substrate × 100). Expression of the *IgASE1* transgene was induced by the addition of galactose to yeast cultures. Only C18 substrates with a double bond at the Δ9 position were elongated by IgASE1. All values are the means of triplicates from three separate experiments.

Table 3
Relative conversion of exogenously supplied LA (C18:2n-6) and ALA (C18:3n-3) by wild-type and mutant IgASE1 expressed in yeast

Mutation	IgASE1 (QAFHH)		GlnHis (Q→H)		GlnAla (Q→A)		GlnPhe (Q→F)		HeloGln (AF→VY)		HeloHis (QAF→HVY)	
Exogenous fatty acids	+18:2	+18:3	+18:2	+18:3	+18:2	+18:3	+18:2	+18:3	+18:2	+18:3	+18:2	+18:3
Conversion (%)	55	48	31	20	42	33	3.3	1	15	8.4	5.4	2.7
% IgASE1	100	100	61	42	76	69	6	2.1	27	17.5	9.8	5.6

A 5-bp deletion in *ELOVL4* is associated with two related forms of autosomal dominant macular dystrophy

Kang Zhang[1,11], Marina Kniazeva[3], Min Han[3], Wen Li[4], Zhengya Yu[1], Zhenglin Yang[1], Yang Li[1], Michael L. Metzker[5], Rando Allikmets[6], Donald J. Zack[1,2], Laura E. Kakuk[7], Pamela S. Lagali[8], Paul W. Wong[8,9,10], Ian M. MacDonald[9,10], Paul A. Sieving[7], David J. Figueroa[4], Christopher P. Austin[4], Robert J. Gould[4], Radha Ayyagari[7] & Konstantin Petrukhin[4]

Stargardt-like macular dystrophy (STGD3, MIM 600110) and autosomal dominant macular dystrophy (adMD) are inherited forms of macular degeneration characterized by decreased visual acuity, macular atrophy and extensive fundus flecks[1–3]. Genetic mapping data suggest that mutations in a single gene may be responsible for both conditions, already known to bear clinical resemblance[1–3]. Here we limit the minimum genetic region for STGD3 and adMD to a 0.6-cM interval by recombination breakpoint mapping and identify a single 5-bp deletion within the protein-coding region of a new retinal photoreceptor-specific gene, *ELOVL4*, in all affected members of STGD3 and adMD families. Bioinformatic analysis of ELOVL4 revealed that it has homology to a group of yeast proteins that function in the biosynthesis of very long chain fatty acids. Our results are therefore the first to implicate the biosynthesis of fatty acids in the pathogenesis of inherited macular degeneration.

We initially studied three independent United States kindreds with autosomal dominant Stargardt-like macular dystrophy characterized by bilateral, symmetric, atrophic lesions in the macula and by the frequent presence of yellow fundus flecks (Fig. 1). A genetic linkage analysis placed the disease gene in all three families to the known *STGD3* locus on chromosome 6q14 (refs. 1,2). We refined the position of the *STGD3* locus by meiotic breakpoint mapping and limited the critical region to an approximately 1.8-cM interval between markers *D6S1622* and *D6S391* (Fig. 2). To further reduce the minimum genetic region, we recruited a five-generation family with STGD3 from Canada[4], as well as an independent five-generation US family with adMD (ref. 3). Genetic linkage and haplotype analyses performed in both families were previously reported[3,4] and confirmed the localization of disease loci to the *STGD3* interval on chromosome 6q14 (refs. 1,2). Comparison of disease haplotypes revealed that all affected members in the five independent families with STGD3 and adMD shared a common founder haplotype between markers *D6S1625* and *D6S1644*, indicating a single, ancestral, disease-specific mutation. A recombination event in individual V-14 from family UM:H785 (ref. 3) at marker *D6S460* identified the centromeric boundary of the critical interval, thus limiting the minimum genetic region to a 0.6-cM interval between markers *D6S460* and *D6S391*. To identify candidate genes in the *D6S460–D6S391* interval, we analyzed the DNA clones derived from the inclusive Sanger Centre contig Chr_6ctg16. BLASTN analysis of DNA sequences from PAC clones covering the region revealed the presence of two candidate genes (*C6orf7* and *BCKDH*), in which we did not find amino-acid changing or splice-site mutations. Realizing that some retina-specific genes might not be represented in the EST database, we broadened our search for candidate genes by comparing the DNA sequence we obtained from PAC clones with GenBank protein entries using a BLASTX algorithm. This comparison revealed two potential exons in PAC dJ92C4 whose DNA sequences, when translated, showed homology with the members of the yeast ELO family, known to be involved in the biosynthesis of fatty acids[5]. Absence of amino-acid changing or splice-site mutations from *C6orf7* and *BCKDH* indicated this new candidate gene, *ELOVL4* (for elongation of very long chain fatty acids), as the most likely candidate for STGD3 and adMD. Complete sequencing of *ELOVL4* cDNA revealed that it encodes a putative protein of 314 amino acids (Fig. 3) with approximately 35% amino acid identity to members of the ELO gene family in yeast (which encode components of the membrane-bound fatty acid elongation system[5]). Several other members of this gene family have been found in rodents[6,7], in which they were shown to be necessary for fatty acid chain elongation[7]. Human ELOVL4 contains all three features typical for members of the ELO family: a hydropathy plot that predicts five transmembrane segments (Fig. 3*b*); a single HXXHH motif identified with fatty acid desaturase and other dioxy iron cluster proteins[8]; and a strong signal (dilysine motif with the lysines at positions –3 and –5 relative to the carboxy terminus) shown to be responsible for the retention of transmembrane proteins in the endoplasmic reticulum[9–11], the known site of biosynthesis of fatty acids with very long chains[12]. We identified a mouse ortholog of human ELOVL4 (92% amino acid identity) after extending a single mouse skin EST (Fig. 3*a*).

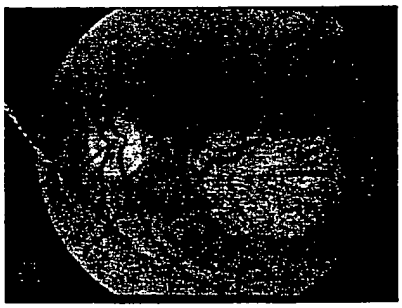


Fig. 1 Fundus photograph from a STGD3 patient with a visual acuity of 20/240 showing a large area of macular atrophy with characteristic yellow flecks surrounding the region of atrophy.

[1]Wilmer Eye Institute and [2]Departments of Molecular Biology & Genetics, and Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [3]Department of Molecular, Cellular and Developmental Biology, Howard Hughes Medical Institute, University of Colorado at Boulder, Boulder, Colorado, USA. Departments of [4]Pharmacology and [5]Human Genetics, Merck Research Laboratories, West Point, Pennsylvania, USA. [6]Departments of Ophthalmology and Pathology, Columbia University, New York, New York, USA. [7]W.K. Kellog Eye Center, University of Michigan, Ann Arbor, Michigan, USA. Departments of [8]Biological Sciences, [9]Ophthalmology, and [10]Medical Genetics, University of Alberta, Edmonton, Alberta, Canada. [11]Present address: Cole Eye Institute, Cleveland Clinic Foundation, Cleveland, Ohio, USA. Correspondence should be addressed to K.P. (e-mail: konstantin_petrukhin@merck.com).

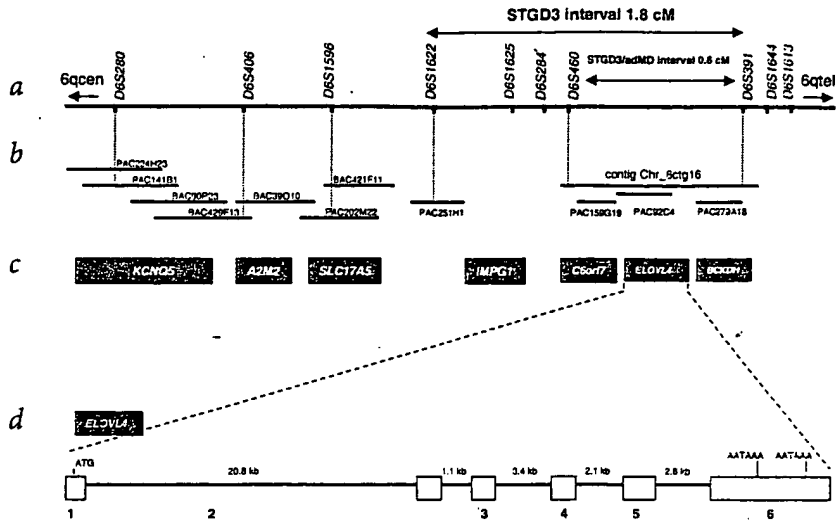
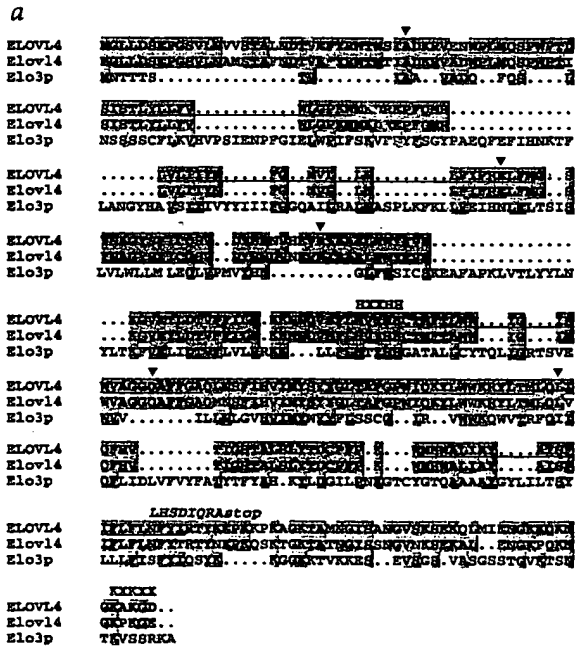


Fig. 2 Physical map of the STGD3/adMD gene region. ***a***, The relative position of the marker loci in the STGD3/adMD gene region. ***b***, A partial contig of overlapping PAC and BAC clones from the disease gene region. Contig Chr6_6ctg16 was downloaded from the Sanger Centre web site. ***c***, Position of seven candidate genes (shown as green rectangles) from the disease gene region which were analyzed for disease-associated mutations at different stages of the positional cloning project. *KCNQ5* (a member of the voltage-gated potassium channel gene family), *A2M2* (a member of the α2-macroglobulin gene family) and *SLC17A5* (a member of the sodium-phosphate cotransporter gene family) are the genes identified by direct sequencing of PAC141B1, BAC90P23, BAC39O10 and PAC202M22 clones. *IMPG1* designates the interphotoreceptor matrix proteoglycan-1 gene previously mapped to the *STGD3* region. *BCKDH* indicates a gene encoding the E1 β-subunit of branched chain α-ketoacid dehydrogenase which maps to PAC dJ279A1 from 6q, but was incorrectly assigned to 6p21–p22 in the literature. *BCKDH* was previously implicated in maple syrup urine type 1B disease (MIM 248611). *C6orf7* designates a four-exon candidate gene residing on PAC dJ159G19 which was assembled after performing a series of 5′-RACE reactions on two overlapping ESTs. ***d***, Exon organization of the STGD3/adMD disease gene (*ELOVL4*). Exons are depicted as rectangles; sizes of introns are indicated above the solid line. The locations of the ATG initiating codon and two alternative polyadenylation signals of *ELOVL4* are also shown. The distance between the genetic markers and sizes of PAC clones, exons and introns are not drawn to scale.

Northern-blot hybridization of an *ELOVL4* cDNA probe to human retina mRNA showed specific and abundant transcription of a 2.9-kb band. We found approximately 90% lower transcription in human brain and no detectable *ELOVL4* mRNA in the ARPE-19 cell line (derived from cells of the retinal pigment epithelium; Fig. 4*a*). RT–PCR analysis (Fig 4*b*,*c*) with gene-specific primers confirmed the retinal specificity of *ELOVL4* mRNA expression. To determine the cell-type specificity of *ELOVL4* expression in the retina, we carried out *in situ* hybridization on rhesus monkey and mouse retinal sections. We detected strong uniform signal in the photoreceptor layer of both species (Fig. 5), particularly in the region corresponding to the photoreceptor inner segments. These results show that in the adult retina *ELOVL4* is expressed exclusively in photoreceptor cells. Given that we observed *ELOVL4* expression in most photoreceptor cells of the rod-dominant monkey and mouse retinas, but also detected its expression in cone photoreceptors (Fig. 5*a*, arrowheads), it is reasonable to conclude that, similar to the recessive Stargardt-disease gene[13–16], *ELOVL4* is expressed in both rod and cone photoreceptors.



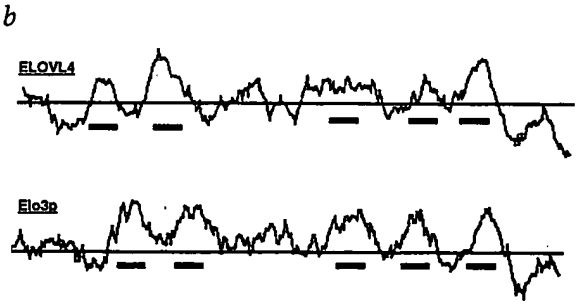


Fig. 3 Comparison of human ELOVL4 and mouse Elovl4 protein sequences with the product of the yeast *ELO3* gene, Elo3p (*a*), and hydropathy profile analysis of the ELOVL4 and Elo3p proteins (*b*). ***a***, Five putative transmembrane domains are underlined. Vertical arrowheads indicate boundaries between exons 1–6 in *ELOVL4*. Conservative and identical residues are indicated in yellow and pink, respectively. Carboxy-terminal dilysine signal responsible for retention in endoplasmic reticulum and the dioxy iron-binding HXXHH motif are indicated above the protein sequence. The truncated C terminus caused by the 5-bp deletion in *ELOVL4* is shown (bold and italics) above the human ELOVL4 sequence. ***b***, Comparison of the ELOVL4 and Elo3p hydropathy profiles. Five regions above the horizontal bars are predicted to span the membrane according to the algorithm of Kyte and Doolittle[30]. *ELOVL1*, *ELOVL2* and *ELOVL3* are yet to be cloned human orthologs of mouse *Elovl1*, *Elovl2* and *Elovl3*, formerly known as *Ssc1*, *Ssc2* and *Cig30*, respectively[6,7]; ELOVL4 exhibits 35% amino acid identity to Elovl1, 40% identity to Elovl2 and Elovl3, and 92% identity to its ortholog Elovl4.

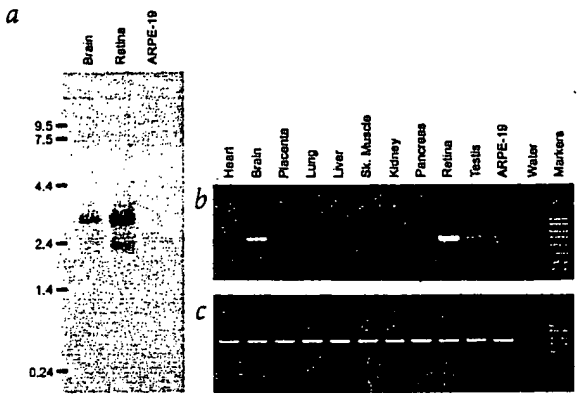


Fig. 4 Expression analysis of *ELOVL4*. **a**, Northern-blot analysis of the disease gene expression. The size of the predominant transcript in the retina (2.9 kb) is consistent with the size of the cloned cDNA. A faint band of 2.0 kb expressed in the retina may represent the product of alternative polyadenylation, as the 3′ UTR of the *ELOVL4* mRNA contains 2 consensus polyadenylation signals (AATAAA) located at positions 2,043 and 2,858, respectively. Hybridization of the cDNA probe to commercially available multiple tissue northern blots confirmed the presence of the 2.9-kb transcript in total brain mRNA, but not in other tissues present on the blot (data not shown). RT–PCR amplification of the disease gene (**b**) and control *GAPDH* fragment (**c**) was carried out on cDNA samples from indicated human tissues and ARPE-19 cells line derived from the cells of the retinal pigment epithelium. After 15 cycles of touchdown PCR, the reaction was allowed to proceed for 15 additional cycles.

We pursued mutation analysis of the candidate gene in four STGD3 kindreds and one family with adMD. We used flanking intronic primers to amplify and sequence all six exons of the candidate gene from several unaffected and heterozygous affected individuals. DNA sequence analysis revealed a 5-bp deletion in exon 6 (Fig. 6*a*) starting at position 797 of the *ELOVL4* cDNA (designated 797–801delAACTT). The deletion results in a frameshift, loss of a fragment of 51 amino acids at the C terminus including a dilysine targeting signal, and synthesis of an aberrant peptide from amino acid 264 to 271, followed by a premature stop codon. We tested for the presence of 797–801delAACTT in all members of five families by direct sequencing. The deletion was present in all affected individuals who inherited the founder haplotype (a total of 76 chromosomes in 5 families) and was absent from unaffected relatives (Fig. 6*a–d*). Consistent with the fact that 797–801delAACTT is not a common polymorphism, we did not find this allele in any of the normal chromosomes from 50 unrelated individuals of predominantly European descent. Moreover, this mutation was absent from 243 controls and 245 AMD individuals constituting a described clinical sample[17].

We were not able to limit the minimum genetic region to an interval entirely contained within *ELOVL4*. Therefore, we cannot exclude the possibility of a disease-causing mutation in an adjacent gene within the 0.6-cM interval; however, linkage disequilibrium between the STGD3/adMD phenotypes and the 797–801delAACTT variant (absent from >1,100 non-STGD3/adMD alleles), photoreceptor specificity of *ELOVL4* expression in the eye and the severe character of the mutation (deletion of the C-terminal portion that includes a putative targeting signal, KXKXX, required for retention of transmembrane proteins in ER; refs. 9–11) indicate that 797–801delAACTT is the disease-causing mutation. To bolster this argument, we analyzed the only other two potential candidate genes found in the *D6S460–D6S391* interval (*C6orf7* and *BCKDH*; Fig. 2), and did not detect amino-acid changing or splice-site mutations in these two genes. Bioinformatic analysis confirmed *ELOVL4* as a candidate gene; based on structural and topological similarities between ELOVL4 and members of the ELO family from yeast[5] and mice[6,7], it seems reasonable to suggest that ELOVL4 represents a photoreceptor-specific component of the fatty acid elongation system residing on the endoplasmic reticulum. Biosynthesis of docosahexaenoic acid (DHA), which contains 22 carbon atoms and represents approximately 50% of the fatty acids in ROS phospholipids[18], requires dietary consumption of the essential α-linolenic acid and a subsequent series of three elongation steps[19,20]. ELOVL4 may be involved in one of the three elongation steps required for DHA biosynthesis. The 797–801delAACTT mutation would likely result in a deficiency in very long chain fatty acids within the retina. Recent data indicate that activity of ABCR (the product of the recessive Stargardt macular dystrophy gene, *ABCA4*, that functions as an outwardly directed flippase for removal of N-retinyliden-phosphatidylethanolamine[21,22]) is strongly dependent on a lipid environment[23], so the role of ELOVL4 in lipid biosynthesis might help to explain the phenotypic similarities between recessive and dominant forms of Stargardt macular dystrophy. The most immediate benefit from isolating the gene responsible for two dominant forms of macular dystrophy is clarification of the pathogenic mechanisms involved in macular degeneration. Although disease genes responsible for four other forms of inherited macular dystrophies have been identified[13,24–26], our results are the first to implicate the biosynthesis of fatty acids in the pathogenesis of at least two related forms of macular degeneration.

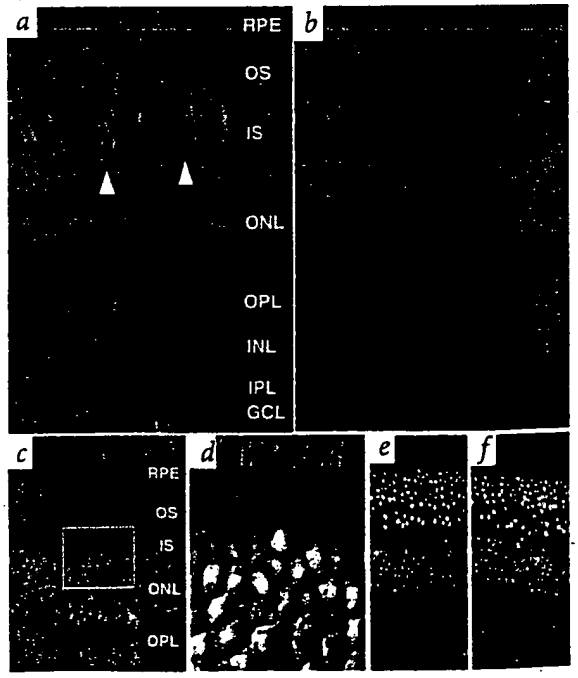


Fig. 5 *In situ* hybridization of *ELOVL4* mRNA in rhesus monkey (*a,b*) and mouse (*c–f*) retina. Specific expression is seen in the photoreceptor layer. **a**, Antisense probe, with signal seen as red fluorescence; retinal cell nuclei are visualized with blue DAPI counterstain; cone photoreceptors are indicated by two arrowheads. **b**, Sense control probe, showing no signal. **c**, Antisense probe on mouse retina; boxed region is seen at higher magnification in (**d**). **e**, Inner mouse retina not visible in (c), demonstrating lack of *Elovl4* mRNA expression in any layer outside the photoreceptor layer. **f**, Sense control probe for (c–e), showing no signal. RPE, retinal pigment epithelium; OS, outer segments of photoreceptors; IS, inner segments of photoreceptors; ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer; GCL, ganglion cell layer.

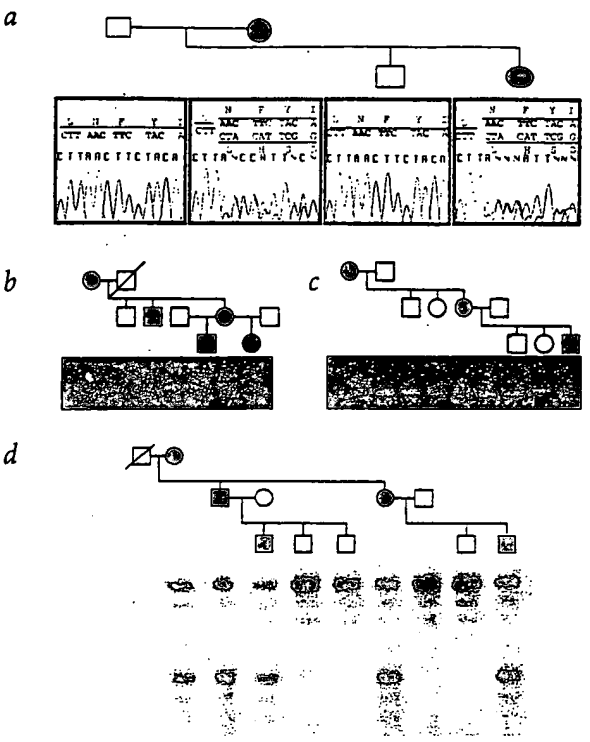


Fig. 6 Segregation of the 797–801delAACT mutation in families with the STGD3 (*a–c*) and adMD (*d*) phenotypes. *a*, Sequence of exon 6 in a small branch of family C. Affected mother and daughter are heterozygous for mutant and wild-type alleles, whereas the unaffected father and son are homozygous for the wild-type allele. The nucleotide sequences and translations are shown above the direct sequencing traces from amplified exon 6. The deletion of five nucleotides (shown in bold) causes a frameshift that results in a premature stop codon. To verify results of heterozygous sequencing, we inserted the PCR products from affected individuals into a cloning vector and confirmed the presence of the 5-bp deletion in 6 of 11 independent clones by sequence analysis. *b–d*, Co-segregation of the 797–801delAACT mutation in branches of families D, A and UM:H785, respectively. PCR products were resolved by running a 4% low gelling temperature agarose gel for 16–20 h (*b,c*) or a 6% acrylamide gel (*d*). The affected individuals who are heterozygous for the 5-bp deletion have 2 PCR bands, whereas normal individuals have only 1 band of 106 bp.

Methods

cDNA cloning and exon/intron organization of *ELOVL4*. We compared DNA sequences of 11 fragments from PAC dJ92C4 generated in Phase 1 High Throughput Genome Sequence Project with GenBank database protein entries using the BLASTX algorithm. This comparison revealed the presence of two potential exons of *ELOVL4* (exons 2 and 4). To identify additional exons that may be located between exons 2 and 4, forward and reverse PCR primers from these exons were synthesized and used to PCR-amplify *ELOVL4* cDNA fragments from human retina "Marathon-ready" cDNA (Clontech). In this RT–PCR experiment, we used a forward primer from exon 2 (5′–GTGTGGAAAATTGGCCTCTG–3′) and a reverse primer from exon 4 (5′–GTCCTCCTGCAACCCAGTTA–3′). The assembly of the DNA sequence results of this PCR product revealed the presence of exon 3. We compared the DNA sequence of exon 4 with the EST database using the BLASTN algorithm to identify additional cDNA sequences. This analysis identified a mouse skin EST (AA791133) with very high degree of similarity to exon 4 of human *ELOVL4*. We compared the DNA sequence of EST AA791133 with the genomic sequence of PAC dJ92C4. Despite the differences between the mouse and human sequences caused by evolutionary divergence, we detected two additional human exons within PAC dJ94C4: exons 5 and 6. To verify the exonic composition of the cDNA, we used forward and reverse PCR primers from exon 2 (5′–GTGTGGAAAATTGGCCTCTG–3′) and exon 6 (5′–GTCAACAACAGTTAAGGCCCA–3′) to PCR-amplify a *ELOVL4* cDNA fragment from human retina "Marathon-ready" cDNA library (Clontech). The DNA sequence of this PCR product provided confirmation of the cDNA sequences that we assembled from *ELOVL4* exons, and corrected the sequencing errors present in the database entry for PAC dJ92C4. To identify the 5′ and 3′ ends of the cDNA, we carried out RACE using the "Marathon-ready" RACE template from human retina. The assembly of the DNA sequence results of these PCR products produced a complete cDNA sequence of the candidate gene (a total of 2,900 bp, including 147 bp of the 5′ UTR followed by a 942-bp ORF). DNA sequencing of the mouse skin cDNA clone (IMAGE 1229776) revealed the sequence of the mouse *Elovl4* cDNA fragment containing exons 4–6. We determined the DNA sequence of the full-length mouse *Elovl4* cDNA after performing a round of 5′-RACE using the "Marathon-ready" RACE template from mouse testis (Clontech).

Genomic sequencing and bioinformatic analysis. We performed DNA sequencing of PAC141B1, BAC90P23, BAC39O10, PAC202M22 and PAC251H1 clones (a total of 497,591 bp of assembled DNA sequence, Fig. 2) as described[24]. We downloaded the DNA sequences of PAC/BAC clones constituting the Chr_6chg16 contig from The Sanger Centre web site (http://www.sanger.ac.uk/HGP/Chr6/) and estimated the size of the minimum genetic region to be within the 400–600-kb range. BLAST analysis of DNA sequences from PAC clones covering the region revealed 14 EST groups. We rejected 12 of 14 EST clusters because they did not contain an identifiable ORF, polyadenylation signal, at least 2 exons, or they were not expressed in the retina. One of the two remaining EST clusters contained a known gene encoding the E1 β-subunit of branched chain α-ketoacid dehydrogenase, *BCKDH*, whereas the second cluster, termed *C6orf7*, contained two overlapping ESTs (retina-specific AA774200 and AI631615), which represented three exons on PAC dJ159G19 with no identifiable homology. We screened *BCKDH* for coding sequence mutations by direct sequencing of its 11 exons and did not find sequence variations in index patients with STGD3. After performing a series of 5′-RACE reactions, we were able to extend the cDNA sequence from the *C6orf7* cluster and assembled a 4-exon candidate gene residing on PAC dJ159G19. We were unable to find amino-acid changing or splice-site mutations in this gene after sequencing the exons amplified from the DNA of index patients with STGD3.

Family material and linkage analysis. Each participant of five European families (A, C, D, STGD3/Canadian and UM:H785) underwent a detailed ophthalmologic examination including a visual acuity measurement and a dilated ophthalmoscopic examination. We made a diagnosis of Stargardt-like macular dystrophy in families A, C and D based on decreased visual acuity and characteristic fundus appearance. Diagnostic criteria and genetic characterization of families UM:H785 and STGD3/Canadian were described[3,4]. Two-point linkage analysis in the 5 families using the marker *D6S460* yielded a maximum lod score of 9.66 (family A), 4.36 (family C), 4.42 (family D) and 3.34 (STGD3/Canadian) at $\theta=0$ and 3.5 at $\theta=0.1$ in UM:H785. The number of family members used in linkage analysis and which were screened for a mutation in *ELOVL4* was (total/affected individuals) 46/26 for family A, 19/12 for family C, 20/11 for family D, 23/11 for STGD3/Canadian family and 39/16 for UM:H785. Family D overlaps with a previously reported 10-generation STGD3 pedigree[2]. Family A represents a described 4-generation pedigree with the reported linkage to chromosome 13q34 ($Z_{max}=4.6$ at $\theta=0.15$ with *D13S71*; ref. 27). Additional clinical studies of children who were initially too young to establish a diagnosis and subsequent genetic linkage ($Z_{max}=9.66$ at $\theta=0$ with *D6S460*) and haplotype analyses placed the disease gene responsible for the Stargardt-like phenotype in this family to the *STGD3* locus on human chromosome 6q14, thus eliminating the locus for autosomal dominant Stargardt-like phenotype on 13q34.

Mutation analysis. We PCR-amplified genomic DNA from STGD3/adMD patients and normal controls using oligonucleotide primers for intronic sequence spaced 30–50 bp from exon/intron junctions. We used the following primer pairs for amplification of *STGD3* exons: exon 1, 5′–TCTCTCTGGGTCTCCGCTT–3′, 5′–GAGGGGAGGCCTTAACATTC–3′; exon 2, 5′–AGCCACTTGCAGGAGTCAGT–3′, 5′–TGATGGTTTTACA

CATTCTCATTTTT–3′; exon 3, 5′–AGCAATCGGAATGCATGAAA–3′, 5′–TTTTCACAGACTGGGGCCTA–3′; exon 4, 5′–TCCATGCCTTGTA CATTTTGTG–3′, 5′–GCGGATCACAAGGTCAAGAG–3′; exon 5, 5′–ACA TCTCAGTGGCTTACTGCCT–3′, 5′–GAAAATTGTCTAAAATGACATTG CAC–3′; exon 6, 5′–GAAGATGCCGATGTTGTTAAAAG–3′, 5′–GTCAA CAACAGTTAAGGCCCA–3′. The primer pair 5′–GCAATATTCCCCAA ATGGATGCACTGG –3′ and 5′–TTGCGGCCGCTTAGGCTCTTTGTATG TCCGAAT–3′ was used to amplify a fragment of exon 6 shown in Fig. 6*b–d*. We amplified DNA (50–100 ng) using *Taq* Gold polymerase (Perkin Elmer Cetus) and a modified touch-down protocol[28] in a total volume of 50 µl. Modified PCR reactions for exon 1 included 10% DMSO. We excised amplification bands from 2% agarose gels, purified them using QIAquick spin columns (Qiagen), and sequenced the DNA using ABI dye-terminator sequencing kits. To assess the degree of polymorphisms in *ELOVL4*, we screened its entire coding sequence by dHPLC analysis in 50 control individuals (100 alleles) of predominantly European descent. Two frequent single-nucleotide polymorphisms (SNPs) were detected: one, in the second intron of *ELOVL4* (IVS2-99T→C), was present in 16% (8/50) of the control sample; the other, 920A→G, resulting in a Met299Val change, was found in 18% (9/50) of controls. Both variants were absent from all 76 disease chromosomes segregating in 5 STGD3/adMD families.

Northern-blot hybridization and RT–PCR. We amplified a 742-bp cDNA fragment from "quick-clone" retinal cDNA (Clontech) with primers 5′–GTGTGGAAAATTGGCCTCTG–3′ and 5′–CATGGCTGTT TTTCCAGCTT–3′, purified it on an agarose gel, and labeled the fragment by random priming with the Rediprime kit (Amersham) in the presence of 3,000 Ci/mmole [α-^{32}P]dCTP (50–100 µCi; Dupont/NEN). We incubated the radiolabeled probe at a concentration of greater than 1×10^6 c.p.m./ml in rapid hybridization buffer (Clontech) for 1 h at 65 °C with human multiple-tissue Northern I (Clontech), and with a custom-made blot containing human retina, brain and ARPE-19 poly(A$^+$) RNA (FRP Grating). We exposed the northern blots to X-ray film or to a Phosphoimager screen for 2–48 h with subsequent quantification of hybridization signals. We performed RT–PCR experiments on "quick-clone" human cDNA samples (Clontech). After 15 cycles of touch-down PCR (ref. 28), we allowed the reaction to proceed for additional 15 cycles.

***In situ* hybridization for *ELOVL4* mRNA.** We ligated a 210-bp human cDNA probe specific for *ELOVL4* (coordinates 561–771 of the ELOVL4 cDNA) to the pCR11-TOPO-Script vector (Stratagene). We linearized the plasmid with either *Hin*dIII or *Not*I to create template DNA. We used SP6 and T7 polymerases with Biotin RNA Labeling Mix (Boehringer) to *in vitro* transcribe antisense and sense riboprobes, respectively. We prepared *in situ* probes specific for the mouse *Elovl4* in the same manner after linearizing the pBluescriptSK(–)-based IMAGE clone 1229776 with *Not*I and *Kpn*I. We carried out *in situ* hybridization as described[24,29].

GenBank accession numbers. *ELOVL4* cDNA, AF277094; mouse *Elovl4* cDNA, AF277093; *ELOVL4* exons 1–6, AF279649, AF279650, AF279651, AF279652, AF279653, AF279654; PAC dJ92C4, AL132875.

Acknowledgments

We thank the members of the sequencing group (Merck Research Labs) for shotgun cloning and sequencing; C.T. Caskey for support and advice; A. Marmorstein and J. Hollyfield for helpful discussions; M. Rayborn for critical reading of the manuscript; J. Seidman, C. Seidman and P. Bither for advice and help in the early phase of this work; Y. Shugart for help with linkage analysis; and patients with STGD3 and adMD and members of their families for participation. Supported by NIH EY00401 and Grant Ritter Fund (K.Z.), Helen Keller Eye Research Foundation (Y.L.), Foundation Fighting Blindness, Hunt Valley, Maryland (D.J.Z., P.A.S., R.A.), E.A. Baker Foundation, AHFMR, and Foundation Fighting Blindness, Canada (P.W.W.), and Howard Hughes Medical Institute, of which M.H. is an assistant investigator. This research was supported in part by funding provided to Johns Hopkins University under a Research Collaboration Agreement between Johns Hopkins University and Merck & Co., Inc.

Received 28 June; accepted 20 October 2000.

1. Stone, E.M. *et al.* Clinical features of a Stargardt-like dominant progressive macular dystrophy with genetic linkage to chromosome 6q. *Arch. Ophthalmol.* **112**, 765–772 (1994).
2. Edwards, A.O. *et al.* Autosomal dominant Stargardt-like macular dystrophy: I. clinical characterization, longitudinal follow-up, and evidence for a common ancestry in families linked to chromosome 6q14. *Am. J. Ophthalmol.* **127**, 426–435 (1999).
3. Griesinger, I.B., Sieving, P.A. & Ayyagari, R. Autosomal dominant macular atrophy at 6q14 excludes CORD7 and MCDR1/PBCRA loci. *Invest. Ophthalmol. Vis. Sci.* **41**, 248–255 (2000).
4. Lagali, P.S. *et al.* Autosomal dominant Stargardt-like macular dystrophy segregating in a large Canadian family. *Can. J. Ophthalmol.* **5**, 315–324 (2000).
5. Oh, C.S., Toke, D.A., Mandala, S. & Martin, C.E. ELO2 and ELO3, homologues of the Saccharomyces cerevisiae ELO1 gene, function in fatty acid elongation and are required for sphingolipid formation. *J. Biol. Chem.* **272**, 17376–17384 (1997).
6. Tvrdik, P. *et al.* Cig30, a mouse member of a novel membrane protein gene family, is involved in the recruitment of brown adipose tissue. *J. Biol. Chem.* **272**, 31738–31746 (1997).
7. Tvrdik, P. *et al.* Role of a new mammalian gene family in the biosynthesis of very long chain fatty acids and sphingolipids. *J. Cell Biol.* **149**, 707–318 (2000).
8. Shanklin, J., Whittle, E. & Fox, B.G. Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* **33**, 12787–12794 (1994).
9. Jackson, M.R., Nilsson, T. & Peterson, P.A. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO J.* **9**, 3153–162 (1990).
10. Jackson, M.R., Nilsson, T. & Peterson, P.A. Retrieval of transmembrane proteins to the endoplasmic reticulum. *J. Cell Biol.* **121**, 317–333 (1993).
11. Schroder, S., Schimmoller, F., Singer-Kruger, B. & Riezman, H. The Golgi-localization of yeast Emp47p depends on its di-lysine motif but is not affected by the ret1-1 mutation in α-COP. *J. Cell Biol.* **131**, 895–912 (1995).
12. Cinti, D.L., Cook, L., Nagi, M.N. & Suneja, S.K. The fatty acid chain elongation system of mammalian endoplasmic reticulum. *Prog. Lipid Res.* **31**, 1–51 (1992).
13. Allikmets, R.L. *et al.* A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy. *Nature Genet.* **15**, 236–246 (1997).
14. Sun, H. & Nathans, J. Stargardt's ABCR is localized to the disc membrane of retinal rod outer segments. *Nature Genet.* **17**, 15–16 (1997).
15. Illing, M., Molday, L.L. & Molday, R.S. The 220-kDa rim protein of retinal rod outer segments is a member of the ABC transporter superfamily. *J. Biol. Chem.* **11**, 10303–10310 (1997).
16. Molday, L.L., Rabin, A.R. & Molday, R.S. ABCR expression in foveal cone photoreceptors and its role in Stargardt macular dystrophy. *Nature Genet.* **25**, 257–258 (2000).
17. Allikmets, R. *et al.* Evaluation of the Best disease gene in patients with age-related macular degeneration and other maculopathies. *Hum. Genet.* **104**, 449–453 (1999).
18. Fliesler, S.J. & Anderson, R.E. Chemistry and metabolism of lipids in the vertebrate retina. *Prog. Lipid Res.* **22**, 79–131 (1983).
19. Sprecher, H., Luthria, D.L. Mohammed, B.S. & Baykousheva, S.P. Reevaluation of the pathways for the biosynthesis of polyunsaturated fatty acids. *J. Lipid Res.* **36**, 2471–2477 (1995).
20. Luthria, D.L., Mohammed, B.S. & Sprecher, H. Regulation of the biosynthesis of 4,7,10,13,16,19-docosahexaenoic acid. *J. Biol. Chem.* **271**, 16020–16025 (1996).
21. Weng, J. *et al.* Insights into the function of Rim protein in photoreceptors and etiology of Stargardt's disease from the phenotype in *abcr* knockout mice. *Cell* **98**, 13–23 (1999).
22. Sun, H., Molday, R.S. & Nathans, J. Retinal stimulates ATP hydrolysis by purified and reconstituted ABCR, the photoreceptor-specific ATP-binding cassette transporter responsible for Stargardt disease. *J. Biol. Chem.* **274**, 8269–8281 (1999).
23. Ahn, J., Wong, J.T. & Molday, R.S. The effect of lipid environment and retinoids on the ATPase activity of ABCR, the photoreceptor ABC transporter responsible for Stargardt macular dystrophy. *J. Biol. Chem.* **275**, 20399–20405 (2000).
24. Petrukhin, K. *et al.* Identification of the gene responsible for Best macular dystrophy. *Nature Genet.* **19**, 241–247 (1998).
25. Stone, E.M. *et al.* A single *EFEMP1* mutation associated with both Malattia Leventinese and Doyne honeycomb retinal dystrophy. *Nature Genet.* **22**, 199–202 (1999).
26. Weber, B.H., Vogt, G., Pruett, R.C., Stohr, H. & Felbor, U. Mutations in the tissue inhibitor of metalloproteinases-3 (TIMP3) in patients with Sorsby's fundus dystrophy. *Nature Genet.* **8**, 352–356 (1994).
27. Zhang, K., Bither, P.P., Park, R., Donoso, L.A., Seidman, J.G. & Seidman, C.E. A dominant Stargardt's macular dystrophy locus maps to chromosome 13q34. *Arch. Ophthalmol.* **112**, 759–764 (1994).
28. Shah, A.B. *et al.* Identification and analysis of mutations in the Wilson disease gene (ATP7B): population frequencies, genotype-phenotype correlation, and functional analyses. *Am. J. Hum. Genet.* **61**, 317–328 (1997).
29. Austin, C.P., Feldman, D., Ida, J. & Cepko, C.L. Vertebrate retinal ganglion cells are selected from competent progenitors by the action of *Notch*. *Development* **121**, 3637–3650 (1995).
30. Kyte, J. & Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).

FEBS 25631

FEBS Letters 510 (2002) 159–165

Identification of a cDNA encoding a novel C18-Δ^9 polyunsaturated fatty acid-specific elongating activity from the docosahexaenoic acid (DHA)-producing microalga, *Isochrysis galbana*[1]

Baoxiu Qi[a], Frédéric Beaudoin[b], Tom Fraser[a], A. Keith Stobart[a,*], Johnathan A. Napier[b], Colin M. Lazarus[a]

[a] *School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK*
[b] *Institute of Arable Crops Research-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS41 9AF, UK*

Received 10 October 2001; revised 8 November 2001; accepted 9 November 2001

First published online 10 December 2001

Edited by Richard Cogdell

Abstract ***Isochrysis galbana*, a marine prymnesiophyte microalga, is rich in long chain polyunsaturated fatty acids such as docosahexaenoic acid (C22:6n-3, $\Delta^{4,7,10,13,16,19}$). We used a polymerase chain reaction-based strategy to isolate a cDNA, designated *IgASE1*, encoding a polyunsaturated fatty acid-elongating activity from *I. galbana*. The coding region of 263 amino acids predicts a protein of 30 kDa that shares only limited homology to animal and fungal proteins with elongating activity. Functional analysis of *IgASE1*, by expression in *Saccharomyces cerevisiae*, was used to determine its activity and substrate specificity. Transformed yeast cells specifically elongated the C18-Δ^9 polyunsaturated fatty acids, linoleic acid (C18:2n-6, $\Delta^{9,12}$) and α-linolenic acid (C18:3n-3, $\Delta^{9,12,15}$), to eicosadienoic acid (C20:2n-6, $\Delta^{11,14}$) and eicosatrienoic acid (C20:3n-3, $\Delta^{11,14,17}$), respectively. To our knowledge this is the first time such an elongating activity has been functionally characterised. The results also suggest that a major route for eicosapentaenoic acid (C20:5n-3, $\Delta^{5,8,11,14,17}$) and docosahexaenoic acid syntheses in *I. galbana* may involve a Δ^8 desaturation pathway.**

Key words: Isochrysis galbana; Polyunsaturated fatty acids (PUFAs); cDNA; Fatty acid elongating activity

1. Introduction

In humans, ω6 (n-6) and ω3 (n-3) long chain polyunsaturated fatty acids (PUFAs) are obtained directly from the diet or synthesised from dietary linoleic acid (LA, C18:2n-6, $\Delta^{9,12}$) and α-linolenic acid (ALA, 18:3n-3, $\Delta^{9,12,15}$), respectively. Long chain PUFAs such as arachidonic acid (AA, C20:4n-6, $\Delta^{5,8,11,14}$), eicosapentaenoic acid (EPA, C20:5n-3, $\Delta^{5,8,11,14,17}$) and docosahexaenoic acid (DHA, C22:6n-3, $\Delta^{4,7,10,13,16,19}$) are considered important as structural components of membrane glycerolipids and as precursors of the eicosanoids, including the biologically active prostaglandins and leukotrienes [1]. AA and DHA are of special importance in the brain and blood vessels, and are considered essential for pre- and post-natal brain and retinal development [2]. Certain human medical conditions may be related to an imbalance in the intake and/or metabolism of the ω3 and ω6 PUFAs resulting in some dysfunction in eicosanoid metabolism [3]. The diet of most modern societies is nowadays relatively low in ω3 PUFAs with a concomitant increased level of ω6 PUFA intake, largely resulting from a preference for plant-seed oils and food products from intensively bred animals. The ω3 PUFAs, EPA and DHA, for example, are usually obtained from fish oils and other marine organisms, and their consumption has decreased significantly in recent years [3]. There is therefore interest in obtaining them from other sources and particularly in genetically engineering the potential to synthesise such products in agronomically important oil seed species.

The conversion of dietary LA and ALA to long chain PUFAs requires further desaturation and elongation reactions, possibly as depicted in Fig. 1. The common route is that both LA and ALA are first desaturated by a Δ^6 desaturase, then a chain-elongating system adds two carbons to the hydrocarbon chains, followed by a further desaturation to yield AA and EPA, respectively [4,5]. These two pathways are referred to here as the ω6 (Δ^6) and ω3 (Δ^6) pathways, since the first step involves Δ^6 desaturation. However, recent evidence suggests that alternative modes of synthesis may exist in some organisms. For example, the production of long chain PUFAs could involve a Δ^8 desaturation step, and a Δ^8 desaturase gene has been isolated from *Euglena* [6]. Such biosynthetic pathways would require the initial elongation of 18:2n-6 and 18:3n-3, rather than desaturation, producing eicosadienoic acid (EDA, C20:2n-6, $\Delta^{11,14}$) and eicosatrienoic acid (ETrA, C20:3n-3, $\Delta^{11,14,17}$), respectively. Δ^8 desaturase activity would then yield dihomo-γ-linolenic acid (DGLA, C20:3n-6, $\Delta^{8,11,14}$) and eicosatetraenoic acid (ETA, C20:4n-3, $\Delta^{8,11,14,17}$) for subsequent conversion to AA and DHA (Fig. 1). To distinguish these modes of synthesis from the previous ones, they are referred to here as the ω6 (Δ^8) and ω3 (Δ^8) pathways.

*Corresponding author. Fax: (44)-117-9257374.
E-mail address: a.k.stobart@bristol.ac.uk (A.K. Stobart).

[1] The nucleotide sequence reported in this paper has been submitted to the GenBank®/EBI Data Bank with accession number AF390174.

Abbreviations: AA, arachidonic acid; ALA, α-linolenic acid; DGLA, dihomo-γ-linolenic acid; DHA, docosahexaenoic acid; EDA, ω6-eicosadienoic acid; EPA, eicosapentaenoic acid; ETA, eicosatetraenoic acid; ETrA, ω3-eicosatrienoic acid; GLA, γ-linolenic acid; LA, linoleic acid; STA, stearidonic acid; GC, gas chromatography; MS, mass spectroscopy; PUFA, polyunsaturated fatty acid

0014-5793/02/$22.00
PII: S0014-5793(01)03247-1

Desaturases involved in PUFA production have been the subject of intensive study in recent years, and many desaturase genes have been isolated [7]. In contrast, only a few genes encoding PUFA-elongating activities have been characterised, and all appear to be related to the *ELO* gene family of *Saccharomyces cerevisiae* [8]. To identify *GLELO*, which encodes an activity involved in elongating the two Δ^6-desaturated C18 PUFAs γ-linolenic acid (GLA, C18:3n-6, $\Delta^{6,9,12}$) and stearidonic acid (STA, C18:4n-3, $\Delta^{6,9,12,15}$), Parker-Barnes et al. [9] constructed a *Mortierella alpina* cDNA library in a yeast expression vector and screened yeast clones for the ability to elongate GLA to DGLA (see Fig. 1). The other genes encoding PUFA-specific elongating activities, *HELO1* from *Homo sapiens* [5] and *F56H11.4* from *Caenorhabditis elegans* [10], were identified via their sequence homology to yeast *ELO2*, which is involved in the elongation of saturated fatty acids [8]. Zhang et al. [11] have recently reported a human retinal photoreceptor-specific gene, *ELOVL4*, which may be involved in one of the three elongation steps required for DHA biosynthesis [12,13]. A 5-bp deletion in the coding region of this gene may be responsible for two related forms of autosomal dominant macular dystrophy [11].

The fatty acid elongation system of both plants and animals is generally considered to comprise a rate-limiting condensing enzyme together with two reductases and one dehydrase; the latter three components are thought to be present constitutively or induced by the condensing enzyme [14,15]. It is therefore considered that it is the condensing enzyme that determines the substrate specificity in terms of chain length and the degree of desaturation of fatty acid substrates. In support of this concept, constitutive expression of the normally seed-specific condensing enzyme encoded by the fatty acid elongation 1 gene (*FAE1*) in transgenic *Arabidopsis* was sufficient to bring about the synthesis of C20 and C22 fatty acids in non-seed tissues [15].

In order for pathways involving Δ^8 desaturation to operate (Fig. 1) it is possible that specific elongases for Δ^9-desaturated PUFAs may exist. Here we describe the isolation and characterisation of a cDNA encoding a protein with such an elongating activity from the DHA-rich microalga, *Isochrysis galbana*.

2. Materials and methods

2.1. Cultivation of I. galbana

I. galbana CCAP 927/1 was obtained from the Culture Collection of Algae and Protozoa, Centre for Coastal and Marine Sciences, Dunstaffnage Marine Laboratory, Oban, Argyll, UK. The algal cultures were grown in f/2 medium [16] in an orbital incubator at 100 rpm and 14°C with continuous white light of intensity of about 30 μmol/m²/s.

2.2. Isolation of total and poly(A)⁺ RNA from I. galbana

Algal cultures were harvested by centrifugation at 3000×*g* for 5 min. The cell pellet was ground to a fine powder under liquid nitrogen with a pestle and mortar before being suspended in 5 ml of TriPure® Isolation Reagent (Roche). Total RNA was then isolated following the manufacturer's protocol. Poly(A)⁺ RNA was prepared with an mRNA isolation kit (Amersham Pharmacia Biotech).

2.3. cDNA library construction

Double-stranded, end-adapted cDNA synthesised using a cDNA synthesis kit (Stratagene) was passed through a Sephacryl S-400 Spun Column (Amersham Pharmacia Biotech) to remove adapters and small cDNA molecules. cDNA eluted from the column was phenol-extracted, ethanol-precipitated and ligated to the arms of the Uni-Zap vector (Stratagene) before packaging into λ phage using the Ready-To-Go Lambda Packaging Kit (Amersham Pharmacia Biotech). A primary library of 1×10^6 pfu was obtained with the majority of the inserts examined ranging from 0.4 to 2 kb. The library was subsequently amplified.

2.4. PCR amplification and cloning of overlapping segments of a putative elongase cDNA

A degenerate primer $EloR_2$ was designed based on the conserved motif MYXYYF/GL in *ELO*-like open reading frames (ORFs) (Fig. 2) and was used in touchdown PCR in conjunction with a universal T3 promoter primer (5′-AATTAACCCTCACTAAAGGG-3′) using an aliquot of the amplified *I. galbana* cDNA library as template. The PCR conditions were: one initial denaturation step of 94°C for 3 min; four cycles of 94°C for 15 s, 52°C for 30 s and 72°C for 45 s; 10 cycles of 94°C for 15 s, 52°C (decreasing by 1°C in each successive cycle) for 30 s, 72°C for 45 s; 25 cycles of 94°C for 15 s, 42°C for 30 s

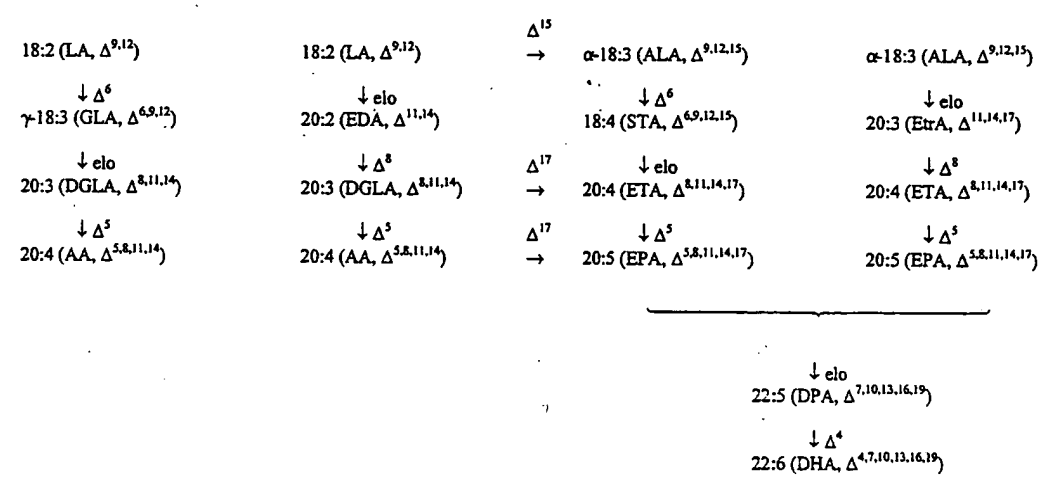


Fig. 1. Possible modes of ω3 and ω6 long chain PUFA biosynthesis. The ω6 (Δ^6) and ω3 (Δ^6) pathways for the synthesis of very long chain PUFAs start with Δ^6 desaturation of the C18 fatty acids, LA and ALA, followed by 2-carbon elongation and then further desaturation and elongation steps. The ω6 (Δ^8) and ω3 (Δ^8) pathways start with chain elongation followed by Δ^8 desaturation, then further desaturation and elongation steps to produce the final products AA and DHA, respectively.

and 72°C for 45 s (increasing by 3 s in each successive cycle); one final extension step of 72°C for 6 min. The resultant putative elongase 5′ end fragment of about 650 bp was gel-purified and cloned into plasmid pCR2.1-TOPO using the TOPO TA cloning system (Invitrogen). The insert of one recombinant plasmid was sequenced and the gene-specific (sense) primer IgEloF$_1$ (5′-ACTCGAAGCTCTTCACATGG-3′) synthesised for use in a further library PCR reaction with a universal M13 forward primer (5′-GTAAAACGACGGCCAGT-3′) under the following conditions: one initial denaturation step of 94°C for 3 min; 10 cycles of 94°C for 15 s, 55°C for 30 s and 72°C for 90 s; 20 cycles of 94°C for 15 s, 55°C for 30 s, and 72°C for 93 s (increasing by 3 s in each successive cycle); one final extension step of 72°C for 6 min. The resultant putative elongase 3′ end fragment of about 850 bp was gel-purified, cloned and sequenced.

2.5. Functional analysis of the IgASE1 ORF by expression in yeast

The entire *IgASE1* coding region was amplified from the *I. galbana* cDNA library with primers IgEloBeg 5′-GGTACCATGGCCCTCG-CAAACGA-3′ (ORF start codon indicated by bold type; underlined sequence is an added *Kpn*I site) and IgEloEnd 5′-TAAGGACATC-CACAATCCAT-3′ (primes 60 bp downstream of the ORF stop codon). The Expand High Fidelity PCR System (Roche) was employed to minimise potential PCR errors. The resultant 860-bp product was cloned in pCR2.1-TOPO (as above) and recombinant plasmids checked for insert orientation. The insert from a selected clone was excised with *Kpn*I and *Sac*I and ligated into the corresponding restriction sites of yeast expression vector pYES2 (Invitrogen), downstream of the *GAL1* promoter. The ligation mixture was used to transform *Escherichia coli* TOP10 cells (Invitrogen), from which the recombinant plasmid designated pY2ASE1 was isolated and used to transform *S. cerevisiae* strain W303-1A to uracil independence by the lithium acetate method [17]. Expression of *IgASE1* was induced by the addition of galactose to 2% (w/v) to cultures grown on raffinose as described [10,18]. After induction, the cultures were grown for 48 h at 22°C in selective medium with individual fatty acid substrates added to 0.1 M and Tergitol Type NP-40 (Sigma) added to 1%.

2.6. Fatty acid analysis

Microalgal or yeast cells were pelleted, washed and dried under a stream of N_2. Total fatty acids were extracted and transmethylated with methanolic HCl. Fatty acid methyl esters were analysed by gas chromatography (GC) using heptadecanoic acid as an internal standard on a 30 m×0.25 mm DB-23 capillary column (J and W Scientific).

2.7. DNA and protein sequence analysis

DNA sequences were determined using an ABI 377 automatic sequencer (Perkin Elmer). The sequence was translated into amino acid sequence using DNAMAN Sequence Analysis Software (Lynnon Biosoft), and the output transferred to BlastP for homology searches.

Table 1
Composition of the major fatty acids of *I. galbana*

Fatty acid	Mol% total
14:0	21
16:0	9.3
16:1n-7	2.6
18:1n-9	18.1
18:2n-6	4.0
18:3n-3	11
18:4n-3	11.5
20:1n-9	8.8
20:5n-3	1.6
22:6n-3	12.0

The microalgae were harvested after 1 month of growth in culture as described in Section 2. Fatty acids were extracted and converted to their methyl esters before analysis by GC. Each peak was identified by the retention time compared to known standards; values represent the average of determinations made from three independent cultures.

3. Results

3.1. Isolation of a cDNA encoding a PUFA-elongating activity

We selected *I. galbana* as the source organism from which to isolate genes involved in EPA and DHA production. Fatty acid analysis of *I. galbana* cultivated as described above showed that it contained some 12% DHA (Table 1). Other major unsaturated fatty acids present were 18:1n-9, Δ^9, 18:3n-3, $\Delta^{9,12,15}$, 18:4n-3, $\Delta^{6,9,12,15}$ and 20:1n-9, Δ^{11}. It is noteworthy that small amounts of 18:2n-6, $\Delta^{9,12}$ and 20:5n-3, $\Delta^{5,8,11,14,17}$ were also present.

The strategy used to isolate a gene encoding a PUFA-elongating activity was to construct and amplify a directional cDNA library and screen it by PCR using degenerate primers based on single conserved motifs and universal primers located in flanking vector sequences. The degenerate primer EloR$_2$ was designed as the reverse complement of the conserved motif MYXYYF/GL in *ELO*-like ORFs [5,9,11, 10,19]. An alignment of the sequences on which the primer design was based is shown in Fig. 2. The primer combination EloR$_2$/T3 amplified a 650-bp product from the library; sequence analysis of the cloned PCR product revealed a continuous ORF extending from a putative start codon to the MYXYYF/GL motif. The 5′ region of the cDNA consisted of a single G upstream of the putative start codon. A further 16 bases of putative 5′ untranslated region were determined by 5′ rapid amplification of cDNA ends (data not shown). However, the data do not exclude the possibility that the first ATG codon of the cDNA encodes an internal methionine residue.

To amplify the 3′ region of the cDNA from the library the gene-specific sense primer IgEloF1 was used with the universal M13 forward primer. A PCR product of about 850 bp was cloned and sequenced; the 5′ and 3′ PCR product sequences overlapped by 235 bp, confirming that they were ultimately derived from a single gene (*IgASE1*). The presence of 17 bases of poly(A) tail indicated that the cDNA was complete at the 3′ end, and that the 3′ untranslated region of the *IgASE1* mRNA was 253 bases.

Assuming the coding region of the cDNA to be full length, *IgASE1* encodes a protein of 263 amino acids with a molecular mass of 30 kDa. The deduced amino acid sequence of IgASE1 aligned with other proteins with elongating activity is shown in Fig. 2. Hydropathy analysis predicts multiple membrane-spanning regions, and dilysine residues at −3 and −6, relative to the C-terminus, probably locate it in the endoplasmic reticulum [20,21]. A further characteristic of all Elo-like proteins described so far is the presence of the histidine box motif HXXHH, which was also found in IgASE1, albeit with glutamine (Q) substituted for the first histidine.

The amino acid sequence of IgASE1 exhibits only limited similarity to the sequences of *ELO*-type genes (Fig. 2), with at best only 27% identity with GLELO from *M. alpina*, which elongates specifically the two Δ6-desaturated C18 PUFAs, GLA and STA [9]. The second highest identity of 26% was found with Ssc1, which, together with Ssc2 (21% identity) and Cig30 (20% identity), is a member of a mouse gene family involved in the elongation of the very long chain fatty acids that become esterified to sphingolipids [22,23]. Ssc1 could replace Elo3 to synthesise C26:0 for sphingolipid synthesis in a yeast mutant, whilst Cig30 restored levels of C20–C24 fatty acids in an Elo2-deficient mutant. IgASE1 shows 26% identity

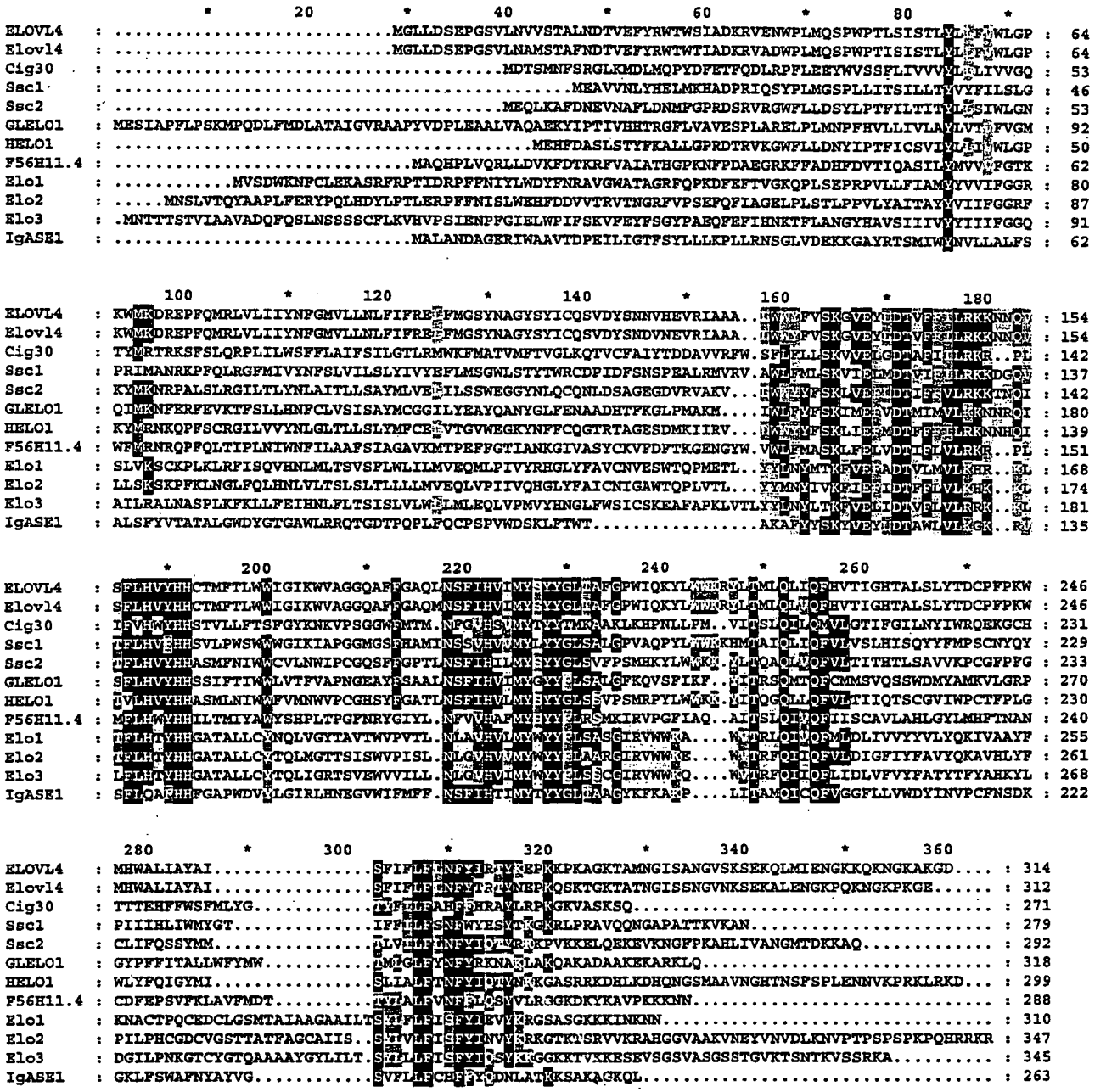


Fig. 2. Alignment of the predicted protein sequence of *I. galbana IgASE1* with sequences predicted from other known elongating activity genes. *ELOVL4* [11] and *HELO1* [5] are from human; *Elovl4* [11], *Cig30*, *Ssc1* and *Ssc2* [22,23] are from mouse; *GLELO1* [9] is from *M. alpina*; *F56H11.4* [10] is from *C. elegans*; *Elo1*, *Elo2* and *Elo3* [8,19] are from yeast.

with human ELOVL4 and 24% identity with the mouse equivalent Elovl4 [11], and 24%, 23% and 21% identity, respectively, with the Elo3, Elo2 and Elo1 proteins from yeast [8]. The lowest identity value of 19% was found with F56H11.4 from *C. elegans* [10]. There is no obvious similarity between IgASE1 and the plant condensing enzyme FAE1, which is required for the synthesis of the mono-unsaturated fatty acid, erucic acid (C22:1n-9, Δ^{13}) in *Arabidopsis* [15].

3.2. *Functional characterisation of IgASE1 in yeast*

Further characterisation of *IgASE1* was achieved by expression in yeast. The complete ORF was amplified from the cDNA library and cloned in pCR2.1-TOPO. Recombinant plasmids were screened for orientation of the insert to allow subsequent excision with the 5′ primer-encoded *KpnI* site and the *SacI* site in the flanking vector sequence. The 5′ *KpnI* site also provided an A base at −3 with respect to the start codon for efficient translation initiation in yeast cells. The *IgASE1* ORF was inserted into the yeast expression vector pYES2 to create pY2ASE1, and this was transferred to the yeast strain, W303-1A.

To determine the substrate specificity of the *IgASE1* prod-

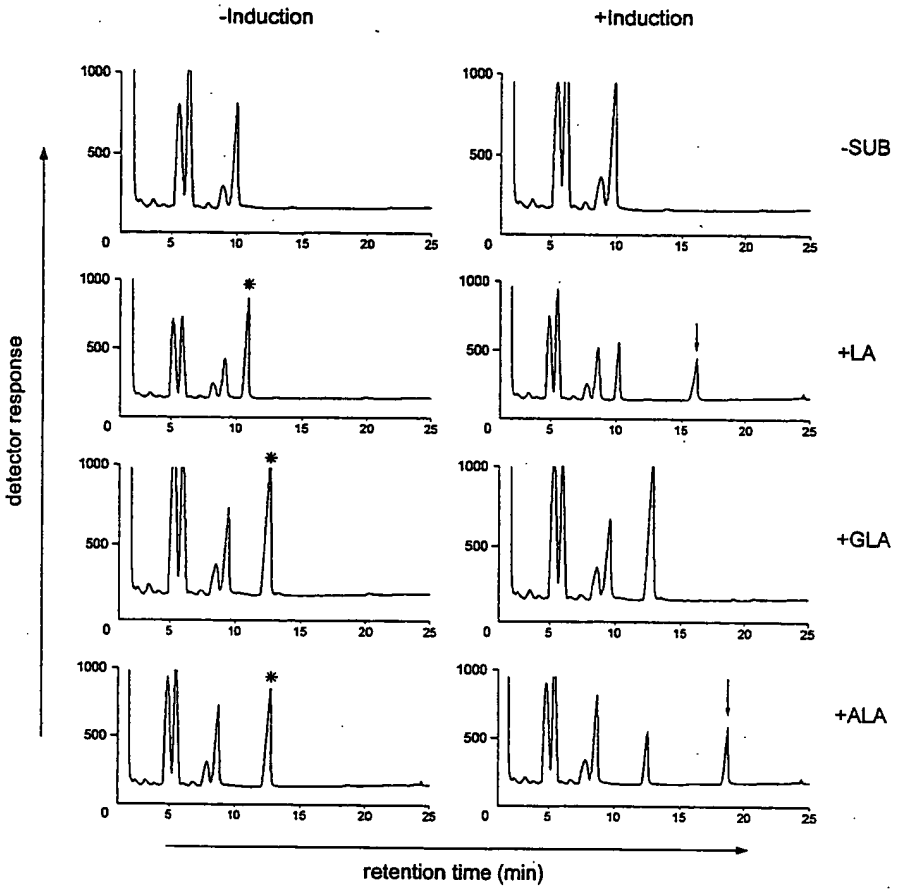


Fig. 3. Gas chromatograms of fatty acid methyl esters extracted from transformed yeast containing pY2ASE1. Yeast cultures were grown in the absence (−SUB) or presence (+LA, +GLA, +ALA) of exogenous fatty acid substrates. Exogenous fatty acids (in the form of their sodium salts) were linoleic acid (18:2n-6, $\Delta^{9,12}$, +LA), γ-linolenic acid (18:3n-6, $\Delta^{6,9,12}$, +GLA) and α-linolenic acid (18:3n-3, $\Delta^{9,12,15}$, +ALA); peaks corresponding to added substrates are indicated by asterisks in the GC traces of uninduced cultures. Expression of *IgASE1* was induced by the addition of galactose to 2%. Yeast cells were harvested after 48 h and fatty acids extracted and assayed as their methyl esters using standard methods. Peaks were identified by co-migration with known standards. Arrows in induced +LA and +ALA traces indicate novel peaks, which were confirmed by GC-MS as EDA (20:2n-6, $\Delta^{11,14}$) and ETrA (20:3n-3, $\Delta^{11,14,17}$), respectively.

uct, the transformed yeast cells were grown in minimal medium with raffinose as the carbon source and supplemented with a range of individual long chain PUFAs. The results (Fig. 3 and Table 2) show that when LA (18:2n-6, $\Delta^{9,12}$) and ALA (18:3n-3, $\Delta^{9,12,15}$) were present in the medium, EDA (20:2n-6, $\Delta^{11,14}$) and ETrA (20:3n-3, $\Delta^{11,14,17}$) accumulated to 9.1 and 8.4 mol% of total fatty acids, respectively. This represented about 45% conversion [product/(product+

Table 2
Fatty acid elongation of different substrates supplied to yeast cells transformed with pY2ASE1

Fatty acid	Mol% of fatty acid							
	−substrate		+LA (18:2n-6)		+ALA (18:3n-3)		+GLA (18:3n-6)	
	+gal	−gal	+gal	−gal	+gal	−gal	+gal	−gal
16:0	28.7	30.2	27.0	28.9	26.6	28.9	30.3	31.0
16:1n-9	41.6	42.4	30.7	25.4	30.1	26.4	24.3	24.6
18:0	6.8	6.1	5.7	5.8	6.3	6.3	6.8	6.2
18:1n-9	22.9	21.3	16.5	13.4	18.4	16.6	14.7	13.4
18:2n-6*	–	–	11.0	26.5	–	–	–	–
18:3n-6*	–	–	–	–	–	–	24.2	24.8
18:3n-3*	–	–	–	–	10.2	21.8	–	–
20:2n-6	–	–	9.1	–	–	–	–	–
20:3n-3	–	–	–	–	8.4	–	–	–
% elongation	0	0	45.3	–	45.2	–	0	–

Exogenous fatty acids supplied as substrates for elongation are indicated by asterisks. The values given are expressed as mol% of total fatty acid methyl esters identified by GC and flame ionization detection. In the case of elongated substrates, this is also expressed as % elongation (product/(product+substrate)×100). Expression of the *IgASE1* transgene was induced by the addition of galactose. Only C18 substrates with a double bond at the Δ^9 position were elongated by the IgASE1. All values are the means of triplicates from three separate experiments.

substrate)×100] of both C18 fatty acid substrates into C20 fatty acids after 48 h growth of the transformed yeast cells. GLA (18:3n-6, $\Delta^{6,9,12}$) was not elongated in the transformed yeast (Fig. 3). Other fatty acids, such as 16:1n-7, Δ^{9}, 18:1n-9, Δ^{9}, 18:4n-3, $\Delta^{6,9,12,15}$, 20:5n-3, $\Delta^{5,8,11,14,17}$ and 22:5n-3, $\Delta^{7,10,13,16,19}$, were also found to be inactive, however, the substrates were always recovered in lipid extracts of the cells (data not shown). Although in these experiments the incorporation of putative fatty acid substrates into membrane components was not investigated, it is noteworthy in the functional characterisation of other fatty acid genes using the yeast system [5,18,25] that the predicted fatty acid conversions have always been observed. These data clearly demonstrate that *IgASE1* encodes a C18-Δ^{9}-specific PUFA-elongating activity.

4. Discussion

We have used PCR in combination with a cDNA library construction to isolate a sequence encoding a C18-Δ^{9}-specific PUFA-elongating activity from the DHA-producing microalga, *I. galbana*. The predicted polypeptide sequence shares only limited identity (up to 27%) with other proteins characterised as having related activity, such as GLELO from *M. alpina*, human HELO1 and the protein encoded by F56H11.4 from *C. elegans*. Other pair-wise alignments of Elo-like proteins (data not shown) produce higher percentage identities, but these probably reflect more the sources of the genes rather than any higher degree of functional relatedness. Hence alignments of the yeast proteins Elo1, Elo2 and Elo3 produce identities in the range 45–56%, while the mammalian proteins HELO1, ELOVL4/Elovl4, Ssc1 and Ssc2 generate identity values in the range 41–57%. The lower identity values arising from comparisons across taxonomic groupings (such as IgASE1 with all the others) result from averaging relatively high degrees of conservation over a limited 'core region' of the proteins, with very little conservation over the large stretches comprising the rest of the sequences. An effect of this is to maintain various consensus motifs identified in known examples of this category of protein, such as KXXEXXDT, HXXMYXYY and TXXQXXQ, while FHXXHH is modified to FH/QXXHH. This last motif is interesting because H/QXXHH also comprises one of the three histidine boxes identified in almost all membrane desaturases and other dioxy iron cluster proteins [7], and deviation from the consensus of HXXHH for the third histidine box to QXXHH is also apparent in various so-called front-end desaturases, including Δ^{5} desaturases from *M. alpina* [18,24] and *C. elegans* [25], Δ^{6} desaturases from *Anabaena* [26] *Borago* [27] and *C. elegans* [28], and the Δ^{8} desaturase from *Euglena gracilis* [6]. Histidine boxes in desaturases and other dioxy iron protein clusters have been implicated in the binding of di-iron [7], but it remains to be determined whether they also play this role in IgASE1 and other proteins with elongating activity.

S. cerevisiae possesses multiple fatty acid elongation systems [8,19], and so by providing the ubiquitous elongase components the yeast expression system has proved a valuable testing ground for characterising the function of putative fatty acid-elongating condensing components from various sources [9,10]. IgASE1 was strictly selective and could only elongate the Δ^{9}-desaturated C18 (ω3 or ω6) PUFAs, LA (18:2n-6, $\Delta^{9,12}$) and ALA (18:3n-3, $\Delta^{9,12,15}$) to yield EDA (20:2n-6, $\Delta^{11,14}$) and ETrA (20:3n-3, $\Delta^{11,14,17}$), respectively. For both fatty acid substrates some 45% was converted to products in the transformed yeast cells.

Some PUFA-elongating activities are active on a range of substrates; human HELO1, for example, can elongate both C18 and C20 to C20 and C22 PUFAs along both the ω3 and ω6 pathways [5]. In contrast, GLELO1 from *M. alpina* showed substrate discrimination, acting specifically on the two products of Δ^{6} desaturation, 18:3n-6 ($\Delta^{6,9,12}$) and 18:4n-3 ($\Delta^{6,9,12,15}$) [9].

It is noteworthy that the gene from *Euglena* encoding a Δ^{8} desaturase activity [6] is specific for the elongation products of IgASE1, ETrA and EDA. This implies that the ω3 (Δ^{8}) desaturase pathway for the synthesis of EPA and DHA is present in *I. galbana* (Fig. 1). It is uncertain, however, whether this coexists with the ω3 (Δ^{6}) desaturase pathway (Fig. 1). Fatty acid analysis of *I. galbana* (Table 1) showed the presence of 18:4n-3, $\Delta^{6,9,12,15}$, a fatty acid known to be specifically synthesised from ALA through the activity of a Δ^{6} desaturase [28–30]. Providing 18:4n-3, $\Delta^{6,9,12,15}$ can be elongated to ETA then this suggests that both ω3 pathways may operate in *I. galbana* for the production of long chain PUFAs.

The fact that the product of *IgASE1* elongates 18:2 to EDA (Table 2 and Fig. 3) and that the *Euglena* Δ^{8} desaturase can act upon this product to produce 20:3n-6, $\Delta^{8,11,14}$ [6] is interesting and implies that the ω6 (Δ^{8}) pathway (Fig. 1) should also operate in *I. galbana*. Although no EDA, GLA or AA could be detected in *I. galbana* it is possible that the enzymes catalysing the synthesis of these intermediates and products are relatively highly active and that there is a rapid flux through these parts of the pathway. Such a flux is also evident in the ω3 pathways in that the intermediates ETrA and ETA were also not detected in fatty acid analyses.

If an ω6 (Δ^{8}) pathway can operate in *I. galbana* then what happens to the products 20:3n-6 and 20:4n-6? Work with *M. alpina* [31] suggested the existence of a desaturase introducing a double bond at the Δ^{17} position of AA to produce EPA, and Spychalla et al. [32] have isolated a *C. elegans* glycerolipid desaturase with ω3 activity on a range of ω6 fatty acid substrates, including DGLA and AA. It is possible that a similar desaturase activity is present in *I. galbana*, rapidly channelling ω6 (Δ^{8}) desaturase pathway products into the ω3 (Δ^{6}) and ω3 (Δ^{8}) pathways for EPA and DHA production (Fig. 1). The presence of 18:4n-3, $\Delta^{6,9,12,15}$ in *I. galbana* implies Δ^{6} desaturase activity (see above). Δ^{6} desaturases are generally non-specific for ω3 and ω6 C18 fatty acids, and can utilise both LA and ALA [29,30]. If this were the case in *I. galbana*, one might anticipate that the ω6 (Δ^{6}) pathway is also active, again with the products rapidly transferred to ω3 fatty acid synthesis via Δ^{17} desaturation (Fig. 1). It is of course possible that the Δ^{6} desaturase is specific for ALA and hence dedicated to ω3 fatty acid production. *I. galbana*, therefore, provides an excellent experimental system for investigating ω3 fatty acid biosynthesis, the interactions between the Δ^{6} and Δ^{8} desaturase pathways and the possible involvement of ω6 metabolism in the production of EPA.

The isolation and characterisation of a novel PUFA elongating activity from *I. galbana*, described in this paper, should make a significant contribution to the determination of structure–function relationships in this class of protein. IgASE1 appears to be less closely related to the Elo-like proteins described hitherto than they are to each other. It is distinct in

having clear specificity for Δ^9-desaturated C18 PUFAs, and this may prove to be of biotechnological as well as biochemical significance.

Acknowledgements: Institute of Arable Crops Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council.

References

[1] Harwood, J.L., Cryer, A., Gurr, M.I. and Dodds, P. (1994) in: The Lipid Handbook (Gunstone, F.D., Harwood, J.H. and Padley, F.B., Eds.), pp. 665–683, Chapman and Hall, London.
[2] Crawford, M.A. (2000) Am. J. Clin. Nutr. 71, 275–284.
[3] Gill, I. and Valivety, R. (1997) Trends Biotechnol. 15, 401–409.
[4] Napier, J.A., Michaelson, L.V. and Stobart, A.K. (1999) Curr. Opin. Plant Biol. 2, 123–127.
[5] Leonard, A.E., Bobik, E.G., Dorado, J., Kroeger, P.E., Chuang, L.T., Thurmond, J.M., Parker-Barnes, J.M., Das, T., Huang, Y.S. and Mukerji, P. (2000) Biochem. J. 350, 765–770.
[6] Wallis, J.G. and Browse, J. (1999) Arch. Biochem. Biophys. 365, 307–316.
[7] Shanklin, J. and Cahoon, E.B. (1998) Annu. Rev. Plant Physiol. Plant Mol. Biol. 49, 611–641.
[8] Toke, D.A. and Martin, C.E. (1996) J. Biol. Chem. 271, 18413–18422.
[9] Parker-Barnes, J.M., Das, T., Bobik, E., Leonard, A.E., Thurmond, J.M., Chuang, L., Huang, Y.S. and Mukerji, P. (2000) Proc. Natl. Acad. Sci. USA 97, 8284–8289.
[10] Beaudoin, F., Michaelson, L.V., Hey, S.J., Lewis, M.J., Shewry, P.R., Sayanova, O. and Napier, J.A. (2000) Proc. Natl. Acad. Sci. USA 97, 6421–6426.
[11] Zhang, K., Kniazeva, K., Han, M., Li, M., Yu, W., Yang, Z., Li, Z., Metzker, Y., Allikmets, R., Zack, D.J., Kakuk, L.E., Lagali, P.S., Wong, P.W., MacDonald, I.M., Sieving, P.A., Figueroa, D.J., Austin, C.P., Gould, R.J., Ayyagari, R. and Petrukhin, K. (2001) Nature Genet. 27, 89–93.
[12] Sprecher, H., Luthria, D.L., Mohammed, B.S. and Baykousheva, S.P. (1995) J. Lipid Res. 36, 2471–2476.
[13] Luthria, D.L., Mohammed, B.S. and Sprecher, H. (1996) J. Biol. Chem. 271, 16020–16025.
[14] Cinti, D.L., Cook, L., Nagi, M.H. and Suneja, S.K. (1992) Prog. Lipid Res. 31, 1–51.
[15] Millar, A. and Kunst, L. (1997) Plant J. 12, 121–131.
[16] Guillard, R.R.L. and Ryther, J.H. (1962) Can. J. Microbiol. 8, 229–239.
[17] Elble, R. (1992) BioTechniques 13, 18–22.
[18] Michaelson, L.V., Lazarus, C.M., Griffiths, G., Napier, J.A. and Stobart, A.K. (1998) J. Biol. Chem. 273, 19055–19059.
[19] Oh, C., Toke, A.D., Mandala, S. and Martin, C.E. (1997) J. Biol. Chem. 272, 17376–17384.
[20] Jackson, M.R., Nilsson, T. and Peterson, P.A. (1990) EMBO J. 9, 3153–3162.
[21] Jackson, M.R., Nilsson, T. and Peterson, P.A. (1993) J. Cell Biol. 121, 317–333.
[22] Tvrdik, P., Asadi, A., Kozak, L.P., Nedergaard, J., Cannon, B. and Jacobsson, A. (1997) J. Biol. Chem. 272, 31738–31746.
[23] Tvrdik, P., Westerberg, R., Silve, S., Asadi, A., Jacobsson, A., Cannon, B., Loison, G. and Jacobsson, A. (2000) J. Cell Biol. 149, 707–717.
[24] Knutzon, D.S., Thurmond, J.M., Huang, Y.S., Chaudhary, S., Bobik, E.G.Jr., Chan, G.M., Kirchner, S.J. and Mukerji, P. (1998) J. Biol. Chem. 273, 29360–29366.
[25] Michaelson, L.V., Napier, J.A., Lewis, M., Griffiths, G., Lazarus, C.M. and Stobart, A.K. (1998) FEBS Lett. 439, 215–218.
[26] Reddy, A.S., Nuccio, M.L., Gross, L.M. and Thomas, T.L. (1993) Plant Mol. Biol. 27, 293–300.
[27] Sayanova, O., Smith, M.A., Lapinskas, P., Stobart, A.K., Dobson, G., Christie, W.W., Shewry, P.R. and Napier, J.A. (1997) Proc. Natl. Acad. Sci. USA 94, 4211–4216.
[28] Griffiths, G., Jackson, F., Brechany, E.Y., Christie, W.W., Stymne, S. and Stobart, A.K. (1996) Phytochemistry 203, 381–386.
[29] Cho, H.P., Nakamura, M.T. and Clarke, S.D. (1999) J. Biol. Chem. 274, 471–477.
[30] Sayanova, O.V., Shewry, P.R. and Napier, J.A. (1999) Plant Physiol. 121, 641–646.
[31] Shimizu, S., Kawashima, H., Akimoto, K., Shinmen, Y. and Yamada, H. (1989) Appl. Microbiol. Biotechnol. 32, 1–4.
[32] Spychalla, J.P., Kinney, A.J. and Browse, J. (1997) Proc. Natl. Acad. Sci. USA 94, 1142–1147.